# Catalytic mechanism for reduction of bicarbonate in the active site of formate dehydrogenase

Master's thesis
Jouni Ruupunen
BioMediTech
University of Tampere
May 2016

**PRO GRADU -TUTKIELMA**

Paikka: TAMPEREEN YLIOPISTO
BioMediTech (BMT)
Tekijä: RUUPUNEN, JOUNI TAPIO
Otsikko: Katalyyttinen mekanismi bikarbonaatin pelkistymiselle formaattidehydrogenaasin aktiivisessa kohdassa
Sivumäärä: 35 sivua, 9 sivua liitteitä
Ohjaajat: Yliopistonlehtori Jarkko Valjakka ja dosentti Ossi Turunen
Tarkastajat: Professori Anne Kallioniemi ja yliopistonlehtori Jarkko Valjakka
Päiväys: Toukokuu 2016

---

## Tiivistelmä

Formaattidehydrogenaasit (FDH) ovat yleensä homodimeerisinä esiintyviä oksidoreduktaaseja, jotka katalysoivat formaatin hapettumista hiilidioksidiksi samalla pelkistäen $NAD^+$:n NADH:ksi. Nämä entsyymit ovat yleisiä metylotrooppisissa hiivoissa ja bakteereissa, joissa niitä tarvitaan metanolin ja metaanin kataboliassa. FDH:t katalysoivat suhteellisen yksinkertaista hydridinsiirtoreaktiota, jossa formaatti luovuttaa hydridin $NAD^+$:n nikotiiniamidirenkaan $C4_N$-atomille. Reaktion tuloksena syntyy hiilidioksidia ja NADH:ia. Viime vuosikymmenten aikana FDH:ja on hyödynnetty koentsyymien uudelleenmuodostusjärjestelmissä, joissa tuotetaan NADH:ia, jonka valmistus tavanomaisilla kemian synteesimenetelmillä olisi kallista. Eräs toinen lupaava FDH:ien käyttökohde on hiilen sitominen ilmakehästä ja meristä. Tähän käytettäisiin käänteisreaktioita, eli hiilidioksidin pelkistämistä formaatiksi. Viimeaikaisten kokeellisten tulosten perusteella voidaan päätellä, että FDH:t voivat pelkistää myös bikarbonaatin formaatiksi.

Tämän pro gradu -tutkielman tavoitteena oli kehittää malli katalyysireaktiolle, jossa bikarbonaatti pelkistyy formaatiksi *Pseudomonas sp*. -bakteerin (kanta 101) FDH:n aktiivisessa kohdassa. Tätä varten tehtiin kolme erilaista mallia klassisia molekyylidynamiikkasimulaatioita varten. PDB-rakenne 2NAD oli kaikkien mallien lähtökohtana. Tärkeimmässä simulaatiossa bikarbonaatti ja NADH asetettiin aktiiviseen kohtaan. Vertailun vuoksi ajettiin myös kaksi muuta simulaatioita, joista ensimmäisessä aktiiviseen kohtaan sijoitettiin formaatti ja $NAD^+$. Tämän simulaation tuloksia verrattiin kirjallisuuteen, jotta muidenkin simulaatioiden tuloksia voitaisiin pitää luotettavina. Toisessa vertailusimulaatiossa tarkasteltiin voisiko hiilidioksidi reagoida takaisin formaatiksi NADH:n vaikutuksesta.

Simulaatiossa bikarbonaatti asettui hydridinsiirtoa suosivaan asentoon. Tähän kului aikaa kuitenkin jopa 3 nanosekuntia. Kokeelliset tulokset olivat samansuuntaisia tämän havainnon kanssa. Bikarbonaatin suotuisan asennon pohjalta luotiin malli pelkistymisreaktiolle. Mallissa NADH:n luovuttama hydridi sitoutuu kovalenttisesti bikarbonaatin hiiliatomiin, mikä synnyttää elektronien ylimäärän. Tästä syystä bikarbonaatin ja hydridin muodostama kompleksi hajoaa formaatti- ja hydroksidi-ioneiksi. Vertailusimulaatiossa formaatti ei asettunut hydridinsiirron kannalta optimaaliseen asentoon; tämä selittyy todennäköisesti klassisen molekyylidynamiikkasimulaation epätarkkuustekijöillä, jotka ovat yleisesti tiedossa. Käänteisreaktio vaikuttaisi olevan mahdollinen, sillä hiilidioksidi kulkeutui ajoittain lähelle NADH:n $C4_N$-atomiin sitoutunutta hydridiä.

**MASTER'S THESIS**

Place: UNIVERSITY OF TAMPERE
BioMediTech (BMT)
Author: RUUPUNEN, JOUNI TAPIO
Title: Catalytic mechanism for reduction of bicarbonate in the active site of formate dehydrogenase
Pages: 35 pages, 9 appendix pages
Supervisors: University Lecturer Jarkko Valjakka and Docent Ossi Turunen
Reviewers: Professor Anne Kallioniemi and University Lecturer Jarkko Valjakka
Date: May 2016

---

## Abstract

Formate dehydrogenases (FDHs) are a set of mostly homodimeric oxidoreductases that catalyze the oxidation of formate to carbon dioxide with the concurrent reduction of $NAD^+$ to NADH. These enzymes are common to methylotropic yeasts and bacteria in which they are vital in the catabolism of methanol and methane. FDHs catalyze relatively simple hydride transfer reaction in which formate donates a hydride to the $C4_N$ atom of the nicotinamide ring of $NAD^+$, yielding carbon dioxide and NADH. During the recent decades, FDHs have been utilized in coenzyme regeneration systems that produce NADH which would be expensive to manufacture by conventional chemical synthesis. Another promising application of FDHs is carbon fixation from the atmosphere and the oceans using the reverse reaction, the reduction of carbon dioxide to formate. Recent experimental evidence suggests that formate dehydrogenases are able to reduce bicarbonate to formate as well.

The aim of this thesis was to develop a model for the catalytic mechanism for the reduction of bicarbonate in the active site of FDH from bacterium *Pseudomonas sp*. strain 101. For this purpose, three different initial models for classical molecular dynamics simulations were constructed. The PDB structure 2NAD was used as a basis for all models. In addition to the main simulation which included bicarbonate and NADH in the active site, two reference simulations were performed as well. In the first one, formate and $NAD^+$ were simulated in the active site to ensure that the results would be comparable to earlier computational studies on the same subject. The purpose of the second reference simulation was to observe if carbon dioxide could react to formate in the presence of NADH.

It was discovered that bicarbonate settled in a conformation ideal for hydride transfer during the simulation. However, the process itself was a time-consuming process, taking up to 3 nanoseconds of simulation time. Experimental results seemed to support this observation. Based on the optimal conformation of bicarbonate, a model for the catalytic mechanism was developed. In the model, the hydride donated by NADH would form a covalent bond to the carbon atom of bicarbonate, resulting in a surplus of electrons within the molecule. This would cause the complex formed by bicarbonate and the hydride to break into formate and hydroxide ions. In the reference simulation with formate the optimal conformations for hydride transfer were not observed, but this was probably due to the inaccuracies of the classical molecular dynamics method described in earlier studies. Carbon dioxide seemed to have a destabilizing effect on the active site. The reverse reaction could still be possible as carbon dioxide occasionally drifted close to the hydride of NADH bound to the $C4_N$.

# Acknowledgements

This thesis project was carried out at BioMediTech, University of Tampere from December 2015 to May 2016. For the most part, I worked together with my supervisor, Jarkko Valjakka, whom I want to thank for introducing me to the fascinating world of structural biology and enzymology. Throughout the twists and turns of the thesis work he was always there to give a little push forward if I had gotten stuck.

I would also like to thank my other supervisor, docent Ossi Turunen of Aalto University, for giving me and Jarkko the opportunity to commence this project in the first place. Ossi's hands-on expertise on formate dehydrogenases which he shared with us through correspondence helped us a lot during the project, providing comprehensive information and novel points of view to our research problems.

My special thanks go to our partners of cooperation at Istanbul Technical University who have conducted various laboratory experiments with formate dehydrogenases. Their experimental results have been vital in evaluating the validity of the computational results.

Last, but definitely not the least, I want to thank my family and all the great friends I was honored to get to know for their unconditional support. The university years have been an unforgettable experience which I will greatly cherish.

Tampere, May 2016

Jouni Ruupunen

# Abbreviations

| | |
|---|---|
| ADH | alcohol dehydrogenase |
| CatDCD | a plugin for manipulating DCD trajectory files |
| CbFDH | formate dehydrogenase from *Candida boindii* |
| CGenFF | CHARMM General Force Field, a force field for small molecules |
| CHARMM | Chemistry at Harvard Macromolecular Mechanics, a force field for MD |
| CmFDH | formate dehydrogenase from *Candida methylica* |
| CSC | Finnish IT Center for Science |
| CtFDH | formate dehydrogenase from *Chaetomium thermophilum* |
| DCD | CHARMM/NAMD binary trajectory file |
| FDH | formate dehydrogenase |
| LDH | lactose dehydrogenase |
| MD | molecular dynamics |
| MDH | malate dehydrogenase |
| NAC | near attack conformation |
| $NAD^+$ | nicotinamide adenine dinucleotide, oxidized |
| NADH | nicotinamide adenine dinucleotide, reduced |
| NAMD | Nanoscale Molecular Dynamics program |
| PDB | Protein Data Bank |
| PseFDH | formate dehydrogenase from *Pseudomonas sp*. strain 101 |
| PSF | Protein Structure File |
| QM/MM | quantum mechanics/molecular mechanics |
| SLURM | Simple Linux Utility for Resource Management |
| TIP3P | a model for simulating water molecules |
| VMD | Visual Molecular Dynamics |

# Table of contents

1 Introduction....1
2 Literature review....3
2.1 Fundamentals of formate dehydrogenases....3
2.2 Structure and localization of formate dehydrogenases....4
2.2.1 Quaternary and tertiary structure....4
2.2.2 Secondary structure....5
2.2.3 Primary structure....7
2.3 Formate dehydrogenase from *Pseudomonas sp*. 101....8
2.3.1 Substrate channel....9
2.3.2 Structure and function of the active site....11
2.3.3 Reaction mechanism....13
2.3.4 Molecular modeling of PseFDH....15
3 Objectives....17
4 Materials and methods....18
4.1 Structure preparation....18
4.2 Simulation phase....19
5 Results and discussion....20
5.1 Overview of the simulations....20
5.2 Conformation of the substrates in the active site....21
5.2.1 Formate and carbon dioxide....21
5.2.2 Bicarbonate....24
5.3 Proposed reaction mechanism for the reduction of bicarbonate....27
5.3.1 Reaction mechanism in detail....27
5.3.2 Kinetic properties of the reaction....28
5.4 Future prospects....30
6 Conclusion....32
7 References....33
Appendix....36

# 1 Introduction

Formate dehydrogenases were discovered in the early 1950s from green pea (*Pisum sativum*) and at that time they were treated as another new dehydrogenases, with no obvious applications. But it was not until the 1970s when the scientific community realized their underlying potential. First off, formate dehydrogenases were used as model enzymes for hydride transfer reactions. The reaction itself is fairly simple. Formate donates a hydride to $NAD^+$ which yields carbon dioxide and NADH. Producing NADH by traditional chemical synthesis is an expensive and laborious process. The idea of NADH regeneration systems utilizing formate dehydrogenases was introduced soon after the first kinetic studies. Such systems were first implemented cost-effectively in the mid-1990s (Tishkov & Popov, 2004).

Meanwhile, the global warming has been progressing at an alarming rate. Governments around the world have been unable to make new meaningful agreements on reducing carbon emissions, let alone to keep their past promises. Atmospheric carbon dioxide is one of the most significant greenhouse gases, and it contributes to the ever-increasing ocean acidification as well. In addition to planting more trees, the idea of enzymatic carbon fixation has emerged during the last decade. This is where formate dehydrogenases come into question. For example, pure enzymes or environmentally harmless microorganisms genetically engineered to produce large amounts of formate dehydrogenase could be placed into the oceans around the world to convert dissolved carbon dioxide and bicarbonate into formate, thus reducing atmospheric carbon and consequently decelerating the global warming.

Carbon fixation with formate dehydrogenases is plausible because the conversion from formate to carbon dioxide and from $NAD^+$ to NADH is essentially reversible, as the major part of all enzymatic reactions. Although, having evolved to formate oxidation, formate dehydrogenases should be studied further and carefully engineered to favor the reverse reaction before using them in the field. An important problem to be solved in this plan is the detoxification of formate produced in the process of carbon fixation.

The unpredictably fast development of computational capabilities of personal computers and computer clusters in the last decades has allowed scientists, including enzymologists and structural biologists, to analyze various enzyme–ligand interactions *in silico*. In this thesis, the methods of molecular modeling and dynamics were used to study the interactions of

bicarbonate, formate and carbon dioxide with either $NAD^+$ or NADH in the active site of formate dehydrogenase from *Pseudomonas sp*. strain 101 which is the most well-known of all formate dehydrogenases. Many detailed computational studies about the behavior of formate in the active site are available, but no similar data on bicarbonate and carbon dioxide have been published. Thus, the aim of this thesis is to provide initial information about the functions of the two latter molecules in the active site and to develop a novel reaction mechanism for the reduction of bicarbonate by the effect of NADH.

# 2 Literature review

## 2.1 Fundamentals of formate dehydrogenases

$NAD^+$-dependent formate dehydrogenases (FDHs) have been the favorite models of enzymologists since the mid-1970s (Popov & Tishkov, 2003). Although the reaction they catalyze is rather simple, chemical and pharmaceutical industries have already utilized formate dehydrogenases in NADH regeneration systems (Tishkov & Popov, 2004). NADH regeneration systems are used in the syntheses of chirally active compounds, such as certain drugs and amino acids, in which the main synthesis reaction consumes NADH. Without constant regeneration, the high-priced NADH required would have to be acquired commercially (Popov & Tishkov, 2003; Tishkov & Popov, 2004). Another yet hypothetical but promising use for formate dehydrogenases is in climate change mitigation. With enzyme engineering, formate dehyrogenases could be made more efficient in utilizing atmospheric $CO_2$ and bicarbonate dissolved in the oceans (Reda *et al*., 2008; Mondal *et al*., 2015).

Vital to methylotropic organisms, formate dehydrogenases are found in many yeasts and aerobic bacteria that use one-carbon compounds as their primary source of carbon. In the catabolism of these compounds, mainly methane and methanol, formate dehydrogenase is the terminal step in the metabolic pathway (Popov & Lamzin, 1994). A schematic representation of the methanol pathway in yeasts is shown in Figure 1. Formate ions are oxidized to carbon dioxide and simultaneously $NAD^+$ is reduced to NADH in a simple hydride transfer reaction where the carbon-hydrogen bond in formate is broken and a new one is formed between the released hydride and $C4_N$ of the nicotinamide ring of $NAD^+$ (Lamzin *et al*., 1994). Water molecules disturb these hydride transfer reactions and in case of most formate dehydrogenases, neither are coordination complexes with metal ions (Lamzin *et al*., 1994). One exception to the rule are particular metal-dependent formate dehydrogenases, which are usually found in bacteria that can operate in anoxic environments, with *E. coli* being the most notable example (Mota *et al*., 2011). However, it should be pointed out that formate dehydrogenases from anaerobic bacteria tend to denaturate in contact with oxygen (Egorov *et al*., 1979). This does not apply to formate dehydrogenases from methylotropic bacteria, as they are stable in air (Popov & Lamzin, 1994; Egorov *et al*., 1979).

The most extensively studied FDH enzyme to date is that of *Pseudomonas sp*., strain 101 (PseFDH), whose residue numbering and structural properties are often used as a reference when studying FDHs from other species. This is because FDHs from bacteria are the stablest known FDHs in general, making them easier to analyze. They have a wide pH optimum, spanning from 6.0 to 9.0. Furthermore, PseFDH is the most thermostable FDH as it does not lose functionality until the temperature of 67.6 ºC is reached (Alekseeva *et al*., 2011; Tishkov & Popov, 2006). Another well-known stable FDH is the one found in yeast *Candida boindii* (Schirwitz *et al*., 2007), which denaturates in 64.4 ºC (Alekseeva *et al*., 2011).

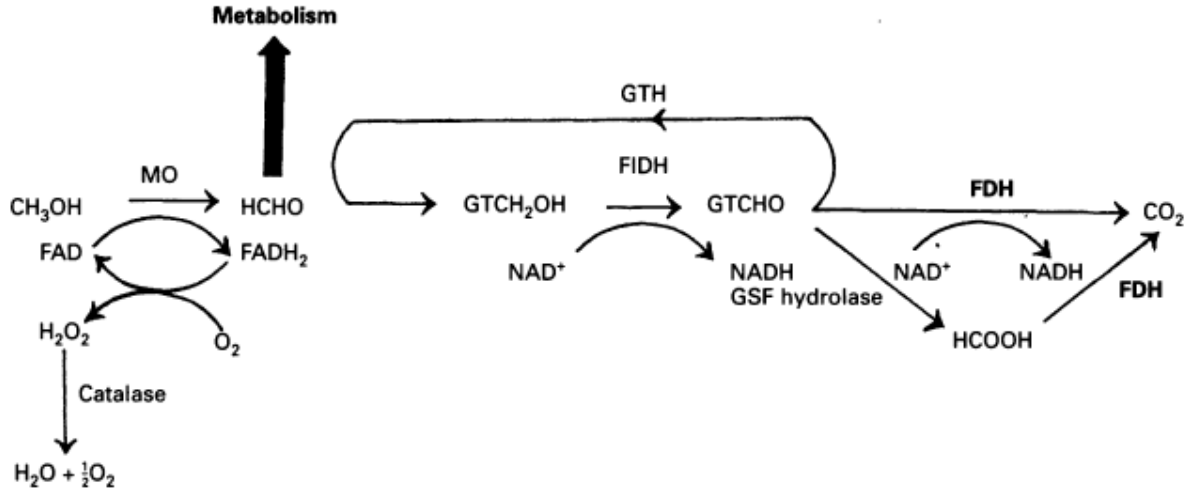


**Figure 1.** The metabolic pathway of methanol catabolism in yeasts. FDH is the terminal step in this pathway as it oxidizes formate anions to $CO_2$. Scheme adapted from Popov & Lamzin (1994).

## 2.2 Structure and localization of formate dehydrogenases

### 2.2.1 Quaternary and tertiary structure

All known formate dehydrogenases found in organisms exist as homodimers. The whole molecule is somewhat globular in shape and each subunit has two Rossmann fold based domains: a coenzyme-binding domain and a catalytic domain (Popov & Lamzin, 1994). The coenzyme-binding domain specifically recognizes and binds $NAD^+$ and the catalytic domain is responsible for catalysis. Interestingly, the two catalytic domains are located on the opposite ends of the FDH enzyme complex, as seen in Figure 2. The domains are linked to each other via single α-helices, designated as αA and α8, respectively (Popov & Lamzin, 1994). *In vivo*,

the two subunits are firmly bound to each other via their coenzyme-binding domains and do not readily dissociate into individual units. However, experimental data has shown that single FDH subunits do have catalytic activity; this was proposed already when the first high resolution structure of PseFDH was determined (Popov & Tishkov, 2003; Lamzin *et al*., 1994). Consequently, conformation changes in one subunit do not affect the other, and thus the subunits function non-cooperatively.

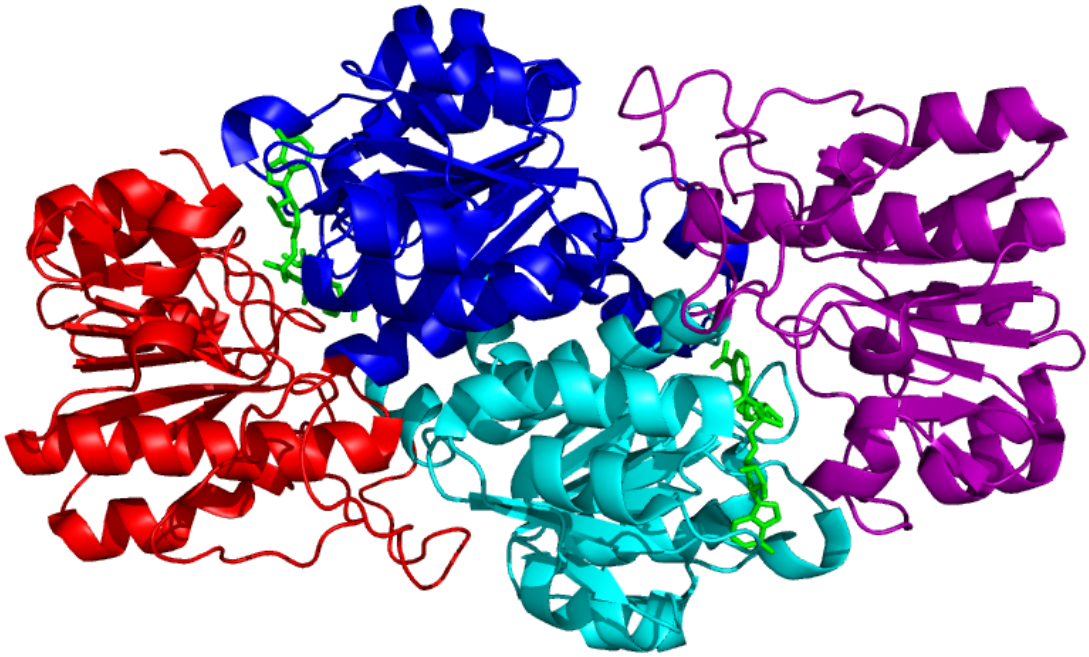

**Figure 2.** Ribbon presentation of the FDH complex with two bound $NAD^+$ molecules (green). Catalytic domains are shown in red and purple, coenzyme binding domains in blue and cyan. Image made with PyMOL (The PyMOL Molecular Graphics System, Version 1.3).

### 2.2.2 Secondary structure

The secondary structure of FDH is very similar compared to those of other dehydrogenase enzymes such as alcohol, lactose and malate dehydrogenases (Popov & Lamzin, 1994). There is a Rossmann fold which is made up by a seven-stranded parallel β-sheet in the coenzyme-binding domain. The β-strands are linked by α-helices. The catalytic domain consists of a

truncated Rossmann fold, with five parallel β-strands encircled by four α-helices (Filippova *et al*., 2005; Popov & Tishkov, 2003; Popov & Lamzin, 1994). Topology diagram of a single subunit of FDH is shown in Figure 3.

When comparing the structures of FDHs from *Pseudomonas sp*. 101 and *Candida boindii* (CbFDH), the main differences lie in the organization of the catalytic domain. CbFDH has a shorter dimerization helix which reduces the contact area between domains (Schirwitz *et al*., 2007). Hence the two domains reside farther away from each other than in the PseFDH subunit. On the contrary, the coenzyme binding domains of both enzymes are nearly identical in shape (Schirwitz *et al*., 2007).

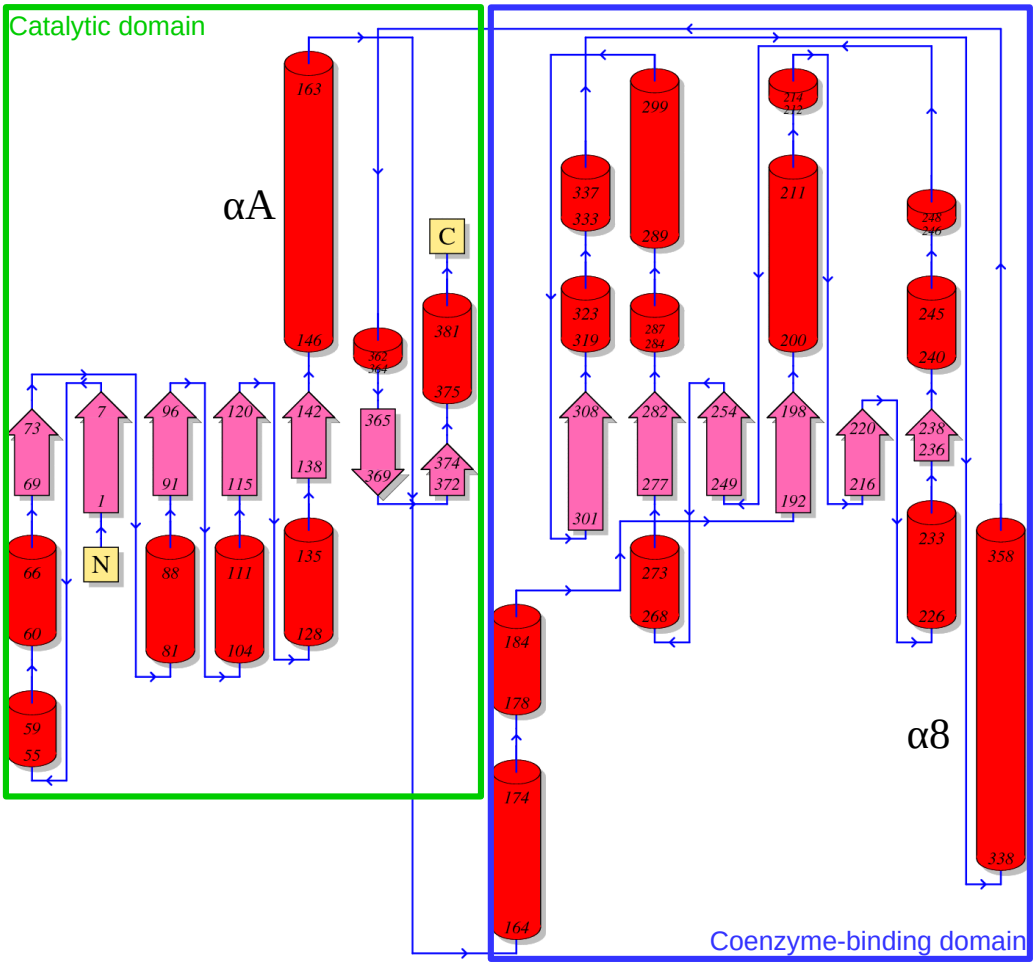


**Figure 3.** Topology diagram of a subunit of PseFDH. α helices are represented by bright red cylinders and β-strands are depicted as pink arrows. The linker helices αA and α8 connected to the other identical subunit are marked in the figure. The catalytic and coenzyme-binding domains are inside green and blue rectangles, respectively. Image generated with the tools of the PDBsum database (Laskowski, 2001).

### 2.2.3 Primary structure

$NAD^+$-dependent formate dehydrogenases have been identified from a diverse set of microorganisms and higher plants. Key characteristics of the most well-known FDHs have been listed in Table 1. Within a taxonomical domain, the sequences of formate dehydrogenases usually show a similarity of 80 percent or more. Especially the amino acids responsible for catalytic activity or binding the substrate and coenzyme are highly conserved. The sequences of amino acid residues delineating the active site are up to 95 percent similar even between bacterial and plant FDHs (Popov & Tishkov, 2003). When comparing whole FDH sequences from different domains, for example bacteria and fungi, they are still approximately 50 percent similar overall (Popov & Tishkov, 2003). An example of sequence alignment of FDHs from different species is shown in Figure 4.

However, sequence alignments of the $NAD^+$-binding region in PseFDH and lactose, malate and alcohol dehydrogenases (LDH, MDH and ADH, respectively) show that these sequences are no more than 20 percent similar (Lamzin *et al*., 1992). Despite the lack of sequence similarity as a whole, some key amino acid residues which form polar contacts with bound $NAD^+$ are considered conserved in all of the above-mentioned enzymes (Lamzin *et al*., 1992).

**Table 1.** Physical and chemical properties of certain $NAD^+$-dependent formate dehydrogenases. Data adapted from Popov & Lamzin (1994).

| **Source** | **Subunits × molecular mass (kDa)** | **Activity (u/mg)** | **pH optimum** | **$K_M^{NAD+}$ (μM)** | **$K_M^{formate}$ (mM)** |
|---|---|---|---|---|---|
| *Pseudomonas sp.* 101 | 2 × 44 | 16.0 | 6.0–9.0 | 110 | 15 |
| *Pseudomonas oxalaticus* | 2 × 100, 2 × 59 | - | 7.5 | 105 | 0.14 |
| *Moxarella sp.* C-1 | 2 × 48 | 6.0 | 6.0–9.0 | 68 | 13 |
| *Candida boindii* | 2 × 36 | 2.4 | 6.5–8.5 | 90 | 13 |
| *Candida methylica* | 2 × 46 | 10.0 | 6.0–9.0 | 100 | 13 |
| *Candida methanolica* | 2 × 43 | 7.5 | 6.5–9.5 | 110 | 3 |
| *Pichia pastoris* NRRL-Y-7556 | 2 × 47 | 8.2 | 6.5–7.5 | 140 | 16 |
| *Pichia pastoris* IFP 206 | 2 × 34 | 2.8 | 7.5 | 270 | 15 |
| *Pisum sativum* | 2 × 42 | 3.7 | 6.0–8.0 | 43 | 1.7 |


**Figure 4.** Sequence alignment of formate dehydrogenases from *Pseudomonas sp*. (FDH_PSESR), *Moxarella sp*. (O08375_MORSP), *Candida boindii* (FDH_CANBO) and *Chaetomium thermophilum* (FDH_CHATD). α-helices are shown in pink, β-strands in pale yellow, amino acids responsible for substrate/coenzyme binding in green and residues vital to catalysis in purple. Alignment made with Clustal Omega at the UniProt website (http://www.uniprot.org/align/).

## 2.3 Formate dehydrogenase from *Pseudomonas sp*. 101

Most of the studies conducted on $NAD^+$-dependent formate dehydrogenases have been focusing on PseFDH. It was the first formate dehydrogenase whose X-ray structure was defined at high resolution (Lamzin *et al*., 1992; Lamzin *et al*., 1994). The apo (without $NAD^+$) and holo (with bound $NAD^+$) structures of PseFDH are available in the RCSB Protein Data Bank, with PDB IDs 2NAC and 2NAD, respectively (see Figure 5). The 2NAD structure has an azide ion bound to the active site in addition to $NAD^+$. This is because azide is a known competitive inhibitor of FDHs and it is similar in shape to $CO_2$ (Lamzin *et al*., 1992). Its binding to the active site causes essentially the same conformational changes in the structure as formate does. Therefore the azide–FDH complex can be used as a model of formate

binding and the transition state of the catalytic reaction (Lamzin *et al*., 1992).

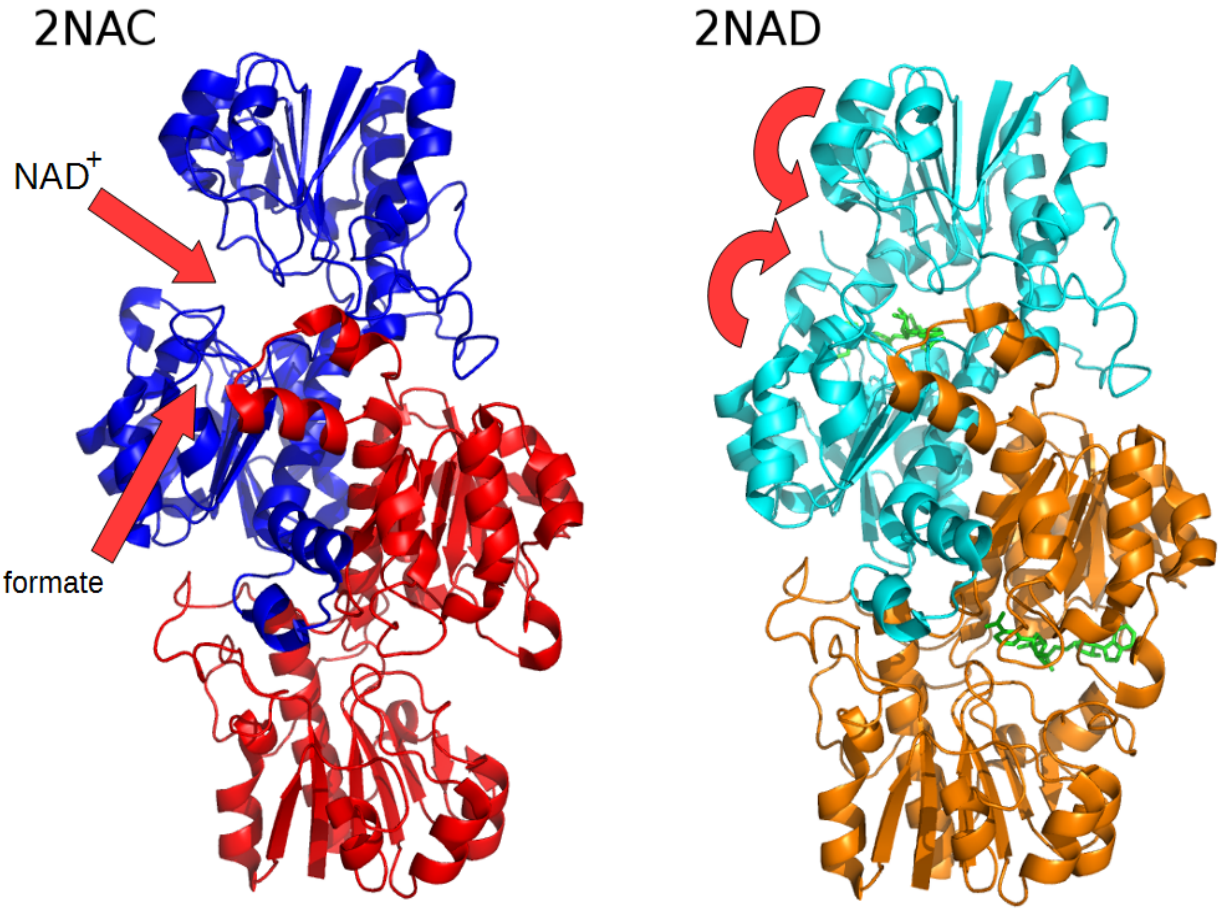


**Figure 5.** Two available PDB structures of PseFDH. 2NAC is the apo form, with chain A colored red and chain B colored blue. 2NAD is the holo form, with bound $NAD^+$ (green) and chains A and B (cyan and orange). Azide and sulphate present in the 2NAD file have been omitted from this representation. In the 2NAC structure, the red arrows denote the entry routes of $NAD^+$ and formate into the active site. $NAD^+$ enters the active site through the substrate channel and formate primarily uses the substrate channel. In the 2NAD structure, the red arrows illustrate the conformational change in the subunit, which results in the blockade of both above-mentioned channels. Image made with PyMOL (The PyMOL Molecular Graphics System, Version 1.3).

### 2.3.1 Substrate channel

The active site of PseFDH is situated inside the subunit, approximately 15 Å from the surface (Popov & Tishkov, 2003). There are two ways which formate can use to enter the active site. Its primary path is the narrow substrate channel. Prior to $NAD^+$ binding, formate can diffuse in via the wide and open coenzyme channel (Popov & Lamzin, 1994). According to molecular dynamics simulations, there are no steric hindrances to formate along this way (Popov &

Tishkov, 2003). Despite the fact that the substrate channel is very narrow, simulations show that formate transport through the channel is slightly more favorable energetically (Popov & Tishkov, 2003). This phenomenon can be accounted for the amino acid residues at the inner end of the channel: Pro97, Arg284 and His332. They have a crucial role in directing formate to a conformation that is favorable to the reaction (Nilov *et al*., 2012) and because of that, they are highly conserved in other FDHs as well (Galkin *et al*., 2002).

Lys286 is the gatekeeper of the substrate channel. Within the pH optimum of PseFDH, 6.0–9.0, its side chain is protonated and thus positively charged because the $pK_a$ of the amino group is 10.53 (http://www.cem.msu.edu/~cem252/sp97/ch24/ch24aa.html; 18/05/2016). Due to its charge, it attracts negatively charged formate and prevents positively charged ions of the bulk solution from entering and possibly blocking the channel (Nilov *et al*., 2012). The channel is quite narrow at both ends, about 5–6 Å in diameter, and therefore only formate analogs similar in size, shape and charge can fit into the channel (Popov & Tishkov, 2003). In the transition from apo to holo form of PseFDH, the conformation changes in the whole subunit cause the substrate channel to close. A loop formed by residues Asn385–Ser390 prevents the entry of solvent molecules into the channel and the inner end of the channel becomes narrower in order to restrain formate in the active site (Popov & Lamzin, 1994). These C terminal residues are not present in the 2NAC structure because Lamzin *et al*. (1994) were unable to determine their correct conformation due to high mobility. They are, however, included in 2NAD since this structure is more tight and stable. The structure of the substrate channel in both apo and holo conformation is displayed in Figure 6.

In the apo form of PseFDH, the coenzyme channel is filled with bulk water. Upon $NAD^+$ binding, water molecules are excluded from the interior parts of the PseFDH subunit. Water molecules in the active site would severely interfere with the hydride transfer reaction, and therefore it has been deduced that the holo structure of PseFDH is too tight for any water molecules (Özgün *et al*., 2015).

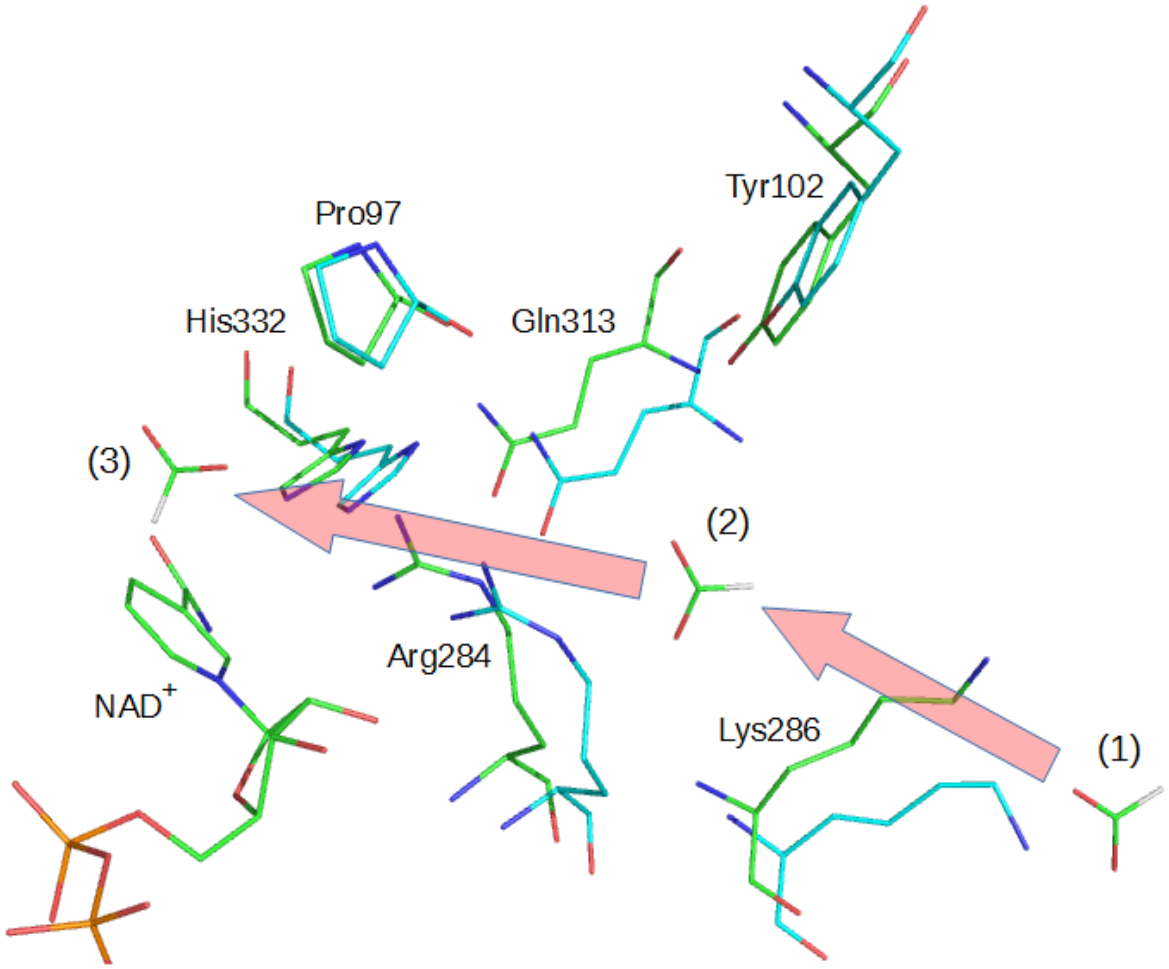


**Figure 6.** Positions of the key residues in both apo (2NAC, cyan carbon backbones) and holo (2NAD, green carbon backbones) conformation in the substrate channel of PseFDH. The transport of formate (yellow carbon atom) to the active site is depicted in three steps: (1) Positively charged Lys286 attracts formate from bulk solution; (2) formate enters the substrate channel and is captured by Arg284 which transports it to the active site (3). Image made with PyMOL (The PyMOL Molecular Graphics System, Version 1.3).

### 2.3.2 Structure and function of the active site

The catalytic domain of PseFDH is formed by amino acid residues 1–146 and 334–393 (Popov & Lamzin, 1994). Key residues responsible for the catalytic activity are Ile122, Asn146, Arg284 and His332 (Nilov *et al*., 2012). All these residues interact with formate via hydrogen bonds during catalysis. Formate is transported to the active site from the substrate channel with the help of Arg284 (Nilov *et al*., 2012). The guanidinium group at the end of the side chain of Arg284 has a $pK_a$ value of 12.48, designating it is strictly protonated under physiological pH (http://www.cem.msu.edu/~cem252/sp97/ch24/ch24aa.html; 18/05/2016).

Inside the active site two another residues form hydrogen bonds to formate. Each of the two oxygen atoms of formate forms two hydrogen bonds. One is connected to Arg284 which has a delocalized positive charge and two amino groups to bond with. The other oxygen atom is bonded to the amino moiety of the carboxamide group of Asn146 and to the nitrogen atom of Ile122 backbone (Nilov *et al*., 2012). Molecular dynamics simulations indicate that the oxygen atoms of formate are held in place by the above-mentioned interactions (Torres *et al*., 1999; Nilov *et al*., 2012). Thus, the carbon–hydrogen bond is free to swing above the $C4_N$ of $NAD^+$ (Torres *et al*., 1999). The hydrophobic amino acid residues Pro97 and Phe98 are located above the nicotinamide ring of $NAD^+$. Being highly conserved in all FDHs, their primary purpose is supposedly to prevent water molecules from entering the active site and disrupting the hydride transfer (Torres *et al*., 1999). Also, their bulky side chains restrict the movement of formate, pushing it towards the $C4_N$ of $NAD^+$ (Popov & Tishkov, 2003). A schematic representation of the active site of PseFDH and the essential polar contacts within is depicted in Figure 7.

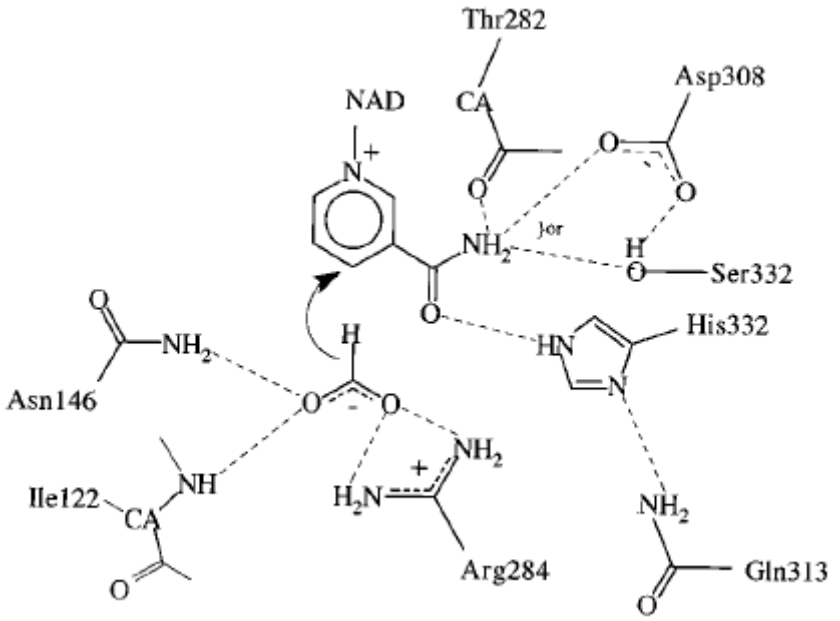


**Figure 7.** The hydrogen bond network in the PseFDH active site with bound formate. Modified scheme originally adapted from Torres *et al*. (1999).

According to some molecular dynamics simulations, His332 can also be involved in formate binding not only during the orientation phase, but during the hydride transfer reaction as well (Nilov *et al*., 2012). This is possible because Gln313 forms a tight, stabilizing hydrogen bond to the imidazole ring of His332 causing the nitrogen atom of His332 facing the active site (often named HE2 in literature) to remain protonated, being able to interact with formate oxygens (Torres *et al*., 1999; Nilov *et al*., 2012). However, the interactions between His332 and formate seem to be only incidental during the hydride transfer (Nilov *et al*., 2012). Instead, His332 has another important role in enabling the reaction. It forms a hydrogen bond to the oxygen atom of the carboxamide group of $NAD^+$ along with Thr282, Asp308 and Ser334; see Figure 7 (Torres *et al*., 1999). Data from simulations has shown that the hydrogen bond between His332 and the carboxamide group is strengthened in tandem with the transition from the apo to holo form of the subunit (Popov & Tishkov, 2003). This shift seems to decrease the volume of the active site, making the reaction between formate and $NAD^+$ more probable (Torres *et al*., 1999). Reflecting the importance of His332, the enzyme activity of PseFDH is lost following its mutation to phenylalanine. The enzyme is still able to bind $NAD^+$ but neither formate nor azide of which the latter is the strongest competitive inhibitor of FDHs (Tishkov *et al*., 1996).

Binding of $NAD^+$ and formate to the active site triggers a conformational change in the PseFDH subunit (see Figure 5). The most notable change is the rotation observed in the catalytic domain. Upon $NAD^+$ binding, the catalytic domain rotates 7.5 degrees in relation to the coenzyme-binding domain about the contact regions of the domains (Popov & Lamzin, 1994). As a result, a novel helix designated as α9 is formed at the C terminus of the subunit. This helix blocks the coenzyme channel, preventing $NAD^+$ from exiting the active site until the hydride transfer. Studies on $NAD^+$ binding suggest that binding of the ADP ribose moiety alone can trigger a significant conformational change towards the holo form in the PseFDH subunit. Nevertheless, the nicotinamide ring and particularly its positive charge are the most important driving forces in the shift from apo to holo (Popov & Lamzin, 1994).

### 2.3.3 Reaction mechanism

The hydride transfer reaction between formate and $NAD^+$ involves breaking a carbon–hydrogen bond and forming a new one (Schiøtt *et al*., 1998). In the reaction, the molecular orbitals of formate and the $C4_N$ of $NAD^+$ overlap and form a short-lived transition state. Next,

the positively charged nicotinamide ring pulls the hydride away from formate. As a result, the $C4_N$ has two hydrogen atoms bound to itself, the nicotinamide ring of newly-formed NADH is neutralized, and formate now oxidized to $CO_2$ starts to diffuse away from the active site (Popov & Lamzin, 1994). This reaction was initially thought to be irreversible (Popov & Tishkov, 2003), but the reverse reaction, the reduction of $CO_2$ to formate, was later found to be possible, but the reaction rate is quite low (Mondal *et al*., 2015), yet improvable with the help of enzyme engineering and the adsorption of FDHs onto electrode surfaces (Sakai *et al*., 2015; Reda *et al*., 2008).

Figuring out the favorable conformations of formate and $NAD^+$ during the hydride transfer reaction has been one of the most important objectives in formate dehydrogenase research. Previously it was thought that linear hydride transfer reactions are energetically favored over any other types (Wu & Houk, 1987). For example, the hydrogen atom of formate would have to point directly towards the $C4_N$ of $NAD^+$ for the hydride transfer to occur. However, an early computational study conducted by Wu and Houk (1987) demonstrated that the transition structures for hydride transfer from methylamine and 1,4-dihydropyridine to methyleniminium are energetically more stable in syn conformation than in anti (see Figure 8). The authors concluded that the syn conformation maximizes the overlap of the molecular orbitals of hydride and the acceptor carbon. The syn conformation is preferred in hydride transfer reactions involving $NAD^+$/NADH as well (Wu *et al*., 1995). Furthermore, the nicotinamide ring of $NAD^+$/NADH has been proven to be planar, both computationally and experimentally (Schiøtt *et al*., 1998; Wu *et al*., 1995).

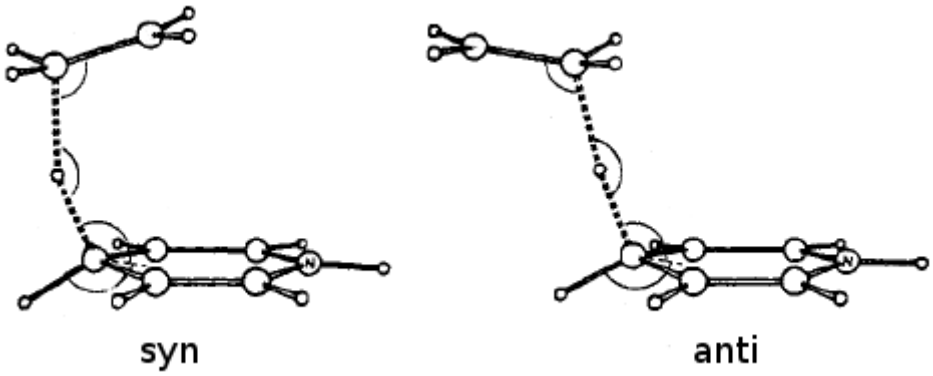


**Figure 8.** The hydride transfer between 1,4-dihydropyridine and methyleniminium in both syn and anti conformations. The syn conformation is energetically more favorable by 5.9 kJ/mol. Modified scheme originally adapted from Wu & Houk (1987).

Torres *et al*. (1999) defined the ideal angle and distance for hydride transfer between the hydrogen atom of formate and the $C4_N$ in their computational study. Designated as the kinetically essential near attack conformation (NAC), the angle C (formate) – H (formate) – $C4_N$ ($NAD^+$) should fall between 132º and 180º, and the H – $C4_N$ distance should be equal or less than 3.0 Å. These parameters are widely used in computational FDH research.

### 2.3.4 Molecular modeling of PseFDH

Molecular modeling and performing molecular dynamics simulations on FDHs were enaled when Lamzin *et al*. (1994) published the first high-resolution X-ray structure of PseFDH. Since it has been possible to compare the results of previous experimental studies to those of computational ones. Molecular dynamics simulations have confirmed the early assumptions about the catalytic mechanisms of FDHs and have provided more information on the roles of certain amino acid residues situated in the active site (Popov & Tishkov, 2003). For example, Torres *et al*. (1999) discovered that the bulky hydrophobic residue Phe98 located above the nicotinamide ring of $NAD^+$ limits the rate of the hydride transfer by frequently blocking the motion of the carbon–hydrogen bond of formate back towards the $C4_N$. Because of this the favorable near attack conformations were perceived only during 1.5 percent of the simulation time.

The most extensive computational study on PseFDH so far was carried out by Nilov *et al*. (2012). They investigated the binding of formate to the active site using three different molecular dynamics methods: classical molecular dynamics, steered molecular dynamics and the most advanced one, QM/MM, which denotes quantum mechanics/molecular mechanics. There were five initial models altogether, and each of them was simulated for 10 nanoseconds. The classical method was used to study formate binding. Calculations on formate transport through the substrate channel were done with the steered method in which formate was placed at the channel entrance and it was steered into the active site with the help of a virtual spring attached to the carbon atom of formate and to the alpha carbon of Asn146. Finally, Nilov *et al*. (2012) selected formate and the nicotinamide ring and the sequential ribose moiety of $NAD^+$ for a more accurate analysis with quantum mechanics/molecular mechanics. Based on the results, the researchers stated that there were vast differences in the probability of formation of the near attack conformation whether the classical or the QM/MM method was used. The carbon–hydrogen bond of formate was oriented towards the $C4_N$ for only 0.5 percent of the time during the classical molecular dynamics simulations. On the contrary, this number was up to 75 percent during one of the QM/MM simulations. The QM/MM calculation method is a very efficient way of analyzing intermolecular interactions in detail. However, the molecules used in these simulations need to be parametrized in greater detail than in conventional molecular dynamics simulations (Senn & Thiel, 2009).

# 3 Objectives

The main objective of this thesis project was to develop models for bicarbonate binding and the concurrent reduction reaction in the active site of formate dehydrogenase from *Pseudomonas sp*. 101 by using the methods of molecular dynamics. Formate dehydrogenases have several practical applications in the field of chemical industry, mainly in NADH regeneration systems, but their ability to reduce $CO_2$ and possibly bicarbonate makes them potential catalysts for carbon fixation from the sea and the atmosphere (Mondal *et al*., 2015). That is, formate dehydrogenases, perhaps after thorough protein engineering, could be used to fight the climate change. In a hypothetical scenario, tons of FDHs would be produced in bacteria and then these enzymes would be poured into large water tanks or even directly into the oceans to react with bicarbonate, preventing its escape to the atmosphere as gaseous $CO_2$.

In an early study on the kinetics of the formate dehydrogenase from *Phaseolus aureus* (mung bean) it was found out that bicarbonate acts as a competitive inhibitor against the conversion from formate to $CO_2$ when $NAD^+$ is bound to the active site (Peacock & Boulter, 1970). Although there is some experimental evidence which implies that FDHs could use bicarbonate as substrate under some circumstances (Sakai *et al*., 2015), no papers about studying this occurrence *in silico* have been published to date. That is, the interactions between bicarbonate and NADH have not been studied in detail. In this study, it was also necessary to compare the binding scheme of bicarbonate to those of formate and $CO_2$, and therefore simulations in which formate and $CO_2$ were docked into the active site were carried out as well.

# 4 Materials and methods

Molecular docking and building of the initial models for molecular dynamics simulations were done with Swiss-PDB Viewer version 4.1 (Guex & Peitsch, 1997). Topology and parameter files for amino acids and all the ligands ($NAD^+$, NADH, formate, bicarbonate, $CO_2$) were taken from the current version (3.0.1) of the CHARMM General Force Field, or CGenFF (Vanommeslaeghe *et al*., 2010). The initial structures were prepared for simulations with the AutoPSF plugin of VMD 1.9.2 (Humphrey *et al*., 1996). The simulations were carried out at the Taito supercluster of the Finnish IT Center for Science (CSC) using NAMD version 2.9 (Phillips *et al*., 2005). Again, VMD was used to visualize and analyze the results of the simulations.

## 4.1 Structure preparation

The PDB structure 2NAD by Lamzin *et al*. (1994) was used as a template for all starting models. In 2NAD, PseFDH is in its native dimeric form. To simplify the upcoming calculation, the A chain was removed from the structure. This action was justified by the fact that the FDH subunits operate independently, without any allosterical regulation (Popov & Tishkov, 2003). The ligands were docked to the active site of the remaining B chain using the coordinates of the azide inhibitor as a reference. After docking, the azide ion was removed from the file as well as the sulphate ion also present in the original structure.

In order to perform molecular dynamics calculations using the CHARMM force field (Brooks *et al*., 1983), protein structure files (PSF) based on the desired structures need to be generated. PSF files contain information about the atoms, types, bonds and partial charges of the residues. A topology file is needed to generate a PSF file, to which the possible interactions described in a parameter file are applied by molecular dynamics programs during calculation (Brooks *et al*., 2009). In this study, the generation of PSF files was automated by AutoPSF, a psfgen-utilizing plugin distributed with the VMD software package. AutoPSF provides a clear graphical user interface for loading topology files to a PDB structure and solvating molecules in TIP3P water with neutralizing $Na^+$ and $Cl^-$ ions included. In each simulation the TIP3P layer was 12 Å thick in each direction and the concentration of NaCl was 0.5 M. The topology files of all ligands are included in the appendices.

After docking, the structures were loaded into VMD. Topology files for ligands were loaded into AutoPSF prior to the generation of the corresponding PSF file. By default, AutoPSF loads the topologies of all amino acids found in proteins on startup. These topology files are from the CGenFF as well. Because atom names in PDB files must be equal to those in topology files, the topology files of the ligands originating from the CGenFF had to be modified, because the atom names and numbering did not match completely. Due to several inconsistencies, every atom was renamed in the topology files of $NAD^+$ and NADH.

The initial structures had $NAD^+$ bound to the active site. It appeared as $NAD^+$ or NADH in the PSF files, depending on which topology file was used in PSF generation. This was possible because the coordinates of hydrogen atoms are not included in the 2NAD structure. Hydrogen atoms are added by the psfgen program based on the information present in the corresponding topology file. Of course, there are differences in bond order and charge between the nicotinamide rings of $NAD^+$ and NADH, but these details have also been defined in the topology files.

## 4.2 Simulation phase

When the PSF files of all three starting models were ready, they were uploaded to the Taito server along with two CHARMM parameter files. One parameter file covered the interactions of amino acid residues and the other contained the corresponding data for the ligands. Also, NAMD configuration files were uploaded to the server. These files were constructed with the help of the NAMDgui plugin of VMD and they retained all the simulation parameters. The same parameters were used in all configuration files. Most importantly, the temperature was set to 298 K, the duration of the energy minimization phase was 2,000 steps (by default, one step is equivalent to one femtosecond) and the maximum duration of the proper simulation phase was adjusted to 10,000,000 steps (that is, 10 nanoseconds).

A SLURM (Simple Linux Utility for Resource Management) batch job file was constructed according to the instructions on the CSC website. This file is used by the supercluster to allocate resources for the computing job submitted. Each simulation was run using 128 cores, with the maximum duration set to 24 hours. The DCD trajectory files produced were up to 50 gigabytes in size and therefore the CatDCD plugin of VMD was used to split the files into smaller and more easily analyzable segments.

# 5 Results and discussion

## 5.1 Overview of the simulations

The three simulations carried out at the Taito supercluster of the Finnish IT Center for Science yielded trajectory files of varying size (Table 2). The bond distances and angles between the transferring hydride and the acceptor carbon were measured with VMD. In all angle measurements the hydride was defined as the vertex and the two rays were set to point at the donor and acceptor carbon atoms. In these simulations, formate carbon and the $C4_N$ of NADH were the donor carbons while the acceptors were the carbon atoms of $CO_2$ and bicarbonate, and the $C4_N$ of $NAD^+$. The results were plotted into two graphs, with one displaying the bond distances and the other showing the angle values (Figures 9 & 10). Another interesting detail worthy of examination was the orientation of the carboxamide group of the nicotinamide ring. This was done by observing the distance between the carboxamide oxygen ($O7_N$) and the hydrogen atom ($H2_N$) bonded to the $C2_N$ (Figure 12). The carboxamide group is free to rotate and may reside in either *cis* or *trans* conformation (Schiøtt *et al*., 1998). Generally, when the oxygen atom of the carboxamide group is oriented towards the active site, the conformation is referred to *cis*. In proportion, the conformation is called *trans* when the oxygen atom is facing the active site.

**Table 2.** Statistics on the simulations performed at CSC. Key numbers of the distances and angles between the hydride and the acceptor carbon are listed in the table as well.

| **Simulation** | 1) formate & $NAD^+$ | 2) $CO_2$ & NADH | 3) bicarbonate & NADH |
|---|---|---|---|
| **Duration (ns)** | 2.70 | 4.05 | 4.80 |
| **Frames** | 54,231 | 81,154 | 96,949 |
| **Min distance (Å)** | 1.91 | 2.41 | 1.89 |
| **Max distance (Å)** | 4.91 | 7.09 | 4.70 |
| **Avg. distance (Å)** | 4.20 | 3.68 | 3.10 |
| **Min angle (º)** | 5.65 | 66.39 | 89.82 |
| **Max angle (º)** | 175.77 | 179.54 | 178.33 |
| **Avg. angle (º)** | 32.97 | 154.19 | 133.82 |

## 5.2 Conformation of the substrates in the active site

### 5.2.1 Formate and carbon dioxide

The range of distances between the $C4_N$ and formate hydrogen and between the hydride bound to $C4_N$ and the carbon atoms of carbon dioxide and bicarbonate showed distinctive differences (see Figure 9). Firstly, formate hydrogen turned away from the $C4_N$ already in the minimization phase of the simulation. The restrictive effect of Phe98 on the movement of formate was observed exactly as described by Torres *et al*. (1999). During the rest of the simulation, the H (formate) – $C4_N$ distance was fixed to 4 Å on average. Thus, the hydride resided too far from the $C4_N$ for hydride transfer to occur. Secondly, $CO_2$ was the most mobile substrate. Within the first nanosecond the carbon atom of $CO_2$ moved back and forth and went as far as 7 Å away from the hydride, momentarily returning to the vicinity of the nicotinamide ring. The transient long distances can be explained by the simultaneous vertical movement of the nicotinamide ring of NADH. Moreover, the simulation with $CO_2$ and NADH was the most unstable one considering the conformation of the active site. The presence of $CO_2$ had a destabilizing effect on the active site, with many key amino acid residues being displaced by the end of the simulation (see Figure 12, panel B3). On the other hand, the destabilization of the PseFDH subunit by $CO_2$ is only natural, because both $CO_2$ and NADH ought to exit the active site, with the subunit returning to the apo conformation.

The drastic swing of the carbon–hydrogen bond of formate in the minimization phase is clearly visible in the graph showing the measures of hydride transfer angles (Figure 10). The initial angle was practically perpendicular at 175.77 degrees and the initial distance from the $C4_N$ was set to 1.90 Å, that is, within the limits of the near attack conformation defined by Torres *et al*. (1999). The average angle between C (formate) – H (formate) – $C4_N$ was no more than 32.97 degrees throughout the simulation. Nevertheless, these results are consistent to earlier studies. It appears that the classical molecular dynamics method has some inaccuracies which could have been avoided by using the QM/MM method instead. QM/MM is better in managing and applying point charges (Nilov *et al*., 2012), but applying the method itself is not trouble-free and could lead to false results if the simulation parameters are not set with caution. Although the simulation with formate was only 2.7 nanoseconds in duration, it is thus highly unlikely that the conformation of formate would have been improved later on.

When it comes to $CO_2$, the angle values seem better than in the simulation with formate (Figure 10). Both the angle values and the distance from the hydride are close to the near attack conformation from 1 to 3.5 nanoseconds of simulation time. After 3.5 nanoseconds the angle begins to decline. This is due to the progressive destabilization of the protein structure. That is, the active site becomes wider as many key amino acid residues which formed hydrogen bonds to $CO_2$ earlier in the simulation drift farther away. Initially the oxygen atoms of $CO_2$ are held in place by the same amino acid residues as the oxygen atoms of formate: Ile122, Asn146 and Arg284. These contacts begin to grow weaker after 3.5 nanoseconds. As the distance between the oxygen atoms of $CO_2$ and the amino groups of the previously mentioned amino acids is extended, the hydrogen bonds become insufficient to coordinate the carbon atom of $CO_2$ towards the hydride, resulting in unorganized spinning of $CO_2$.

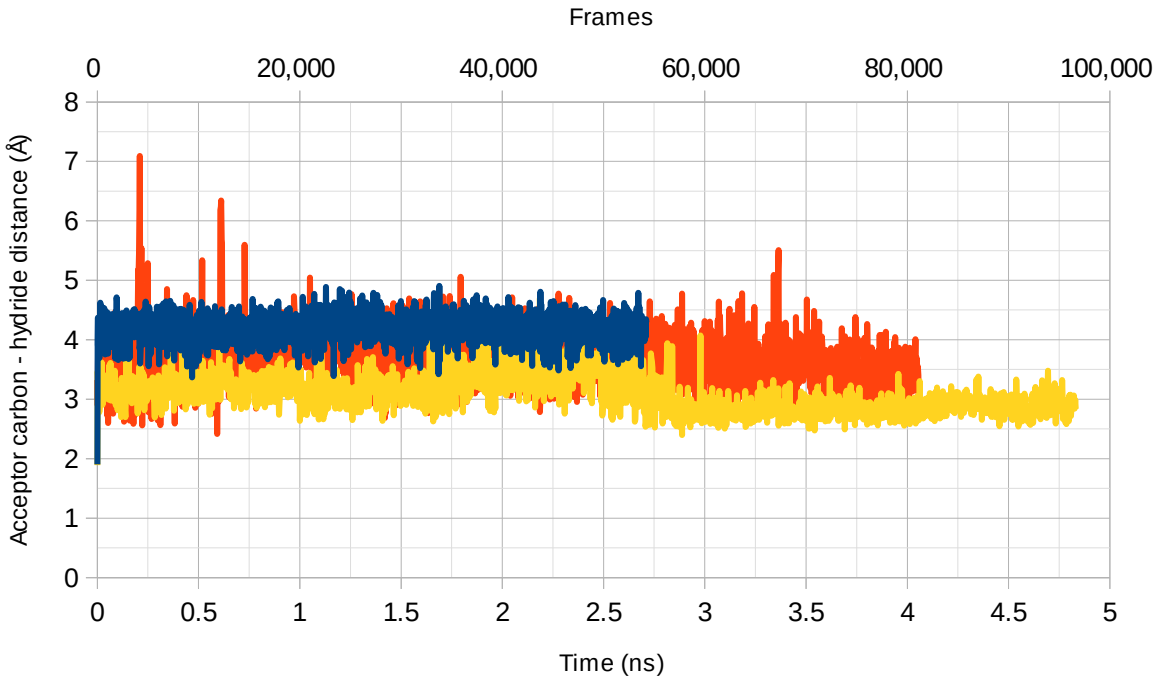


**Figure 9.** Distances between the transferring hydride and its acceptor carbon atom during the course of the three simulations. Simulations with formate, $CO_2$ and bicarbonate are represented by blue, red and yellow lines, respectively.

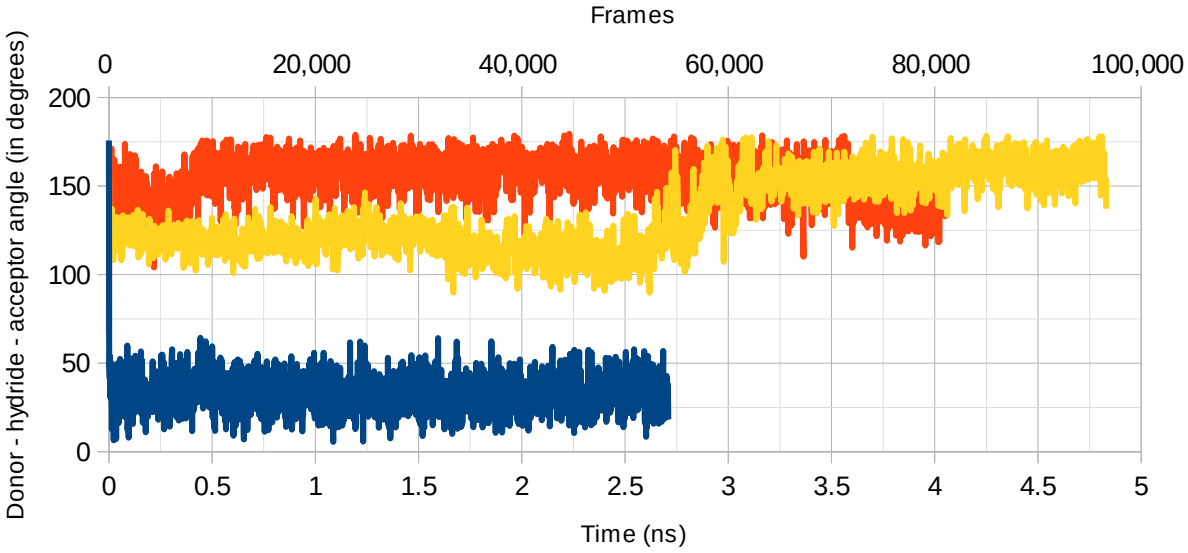


**Figure 10.** Angles between the donor and acceptor carbon atoms of either substrate or $NAD^+$/NADH. The hydride was set as the vertex in angle measurements. Simulations with formate, $CO_2$ and bicarbonate are represented by blue, red and yellow lines, respectively.

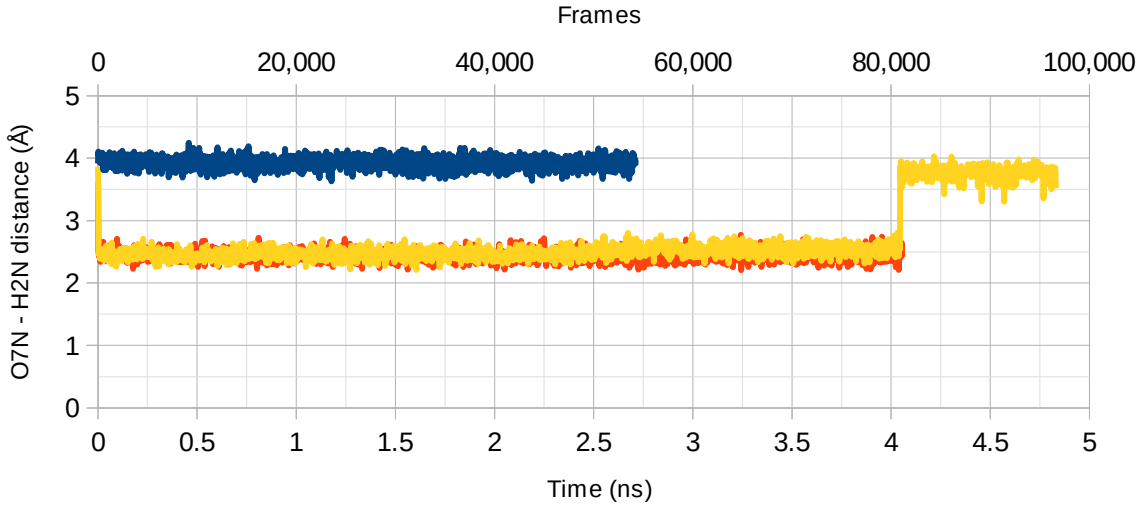


**Figure 11.** Orientation of the carboxamide group of $NAD^+$/NADH during simulations. When the distance between $O7_N$ (carboxamide oxygen) and $H2_N$ (a hydrogen of the nicotinamide ring) is roughly 4 Å, the orientation is *trans*. When it is about 2.5 Å, the orientation is *cis*.

### 5.2.2 Bicarbonate

The initial conformation of bicarbonate was chosen considering the optimality of the hydrogen bond network. Therefore, bicarbonate was placed close to the nicotinamide ring of NADH, in order that the hydroxyl group of bicarbonate was facing the carboxamide group (Figure 13, panel 1). It was found out that the hydrogen bonds between bicarbonate and the residues surrounding it are maximized when the carboxamide group of NADH is rotated into *cis* conformation in which the amino group is pointing towards the active site. For instance, the oxygen atom of the hydroxyl group of bicarbonate could form two hydrogen bonds with the amino group. However, no changes in the conformation of the carboxamide group were made in the 2NAD PDB file which was used as the starting point of the actual simulation. Because the carboxamide group is free to rotate (Schiøtt *et al*., 1998), its orientation should transform from the native *trans* to *cis* if the latter is more energetically favorable.

Compared to formate and $CO_2$, bicarbonate remained relatively close to the nicotinamide ring during the simulation. The maximum distance between the carbon atom of bicarbonate and the hydride of NADH was 4.70 Å and the average distance only 3.10 Å (Figure 9). Shortly after minimization bicarbonate began drifting away from the hydride while rolling over the nicotinamide ring (Figure 13, panels 2 and 3). This change was facilitated by the interactions between bicarbonate oxygens and the amino groups of Asn146 and Arg284. The great shift in angle from roughly 110 degrees to values apt for the near attack conformation between 2.5 and 3.0 nanoseconds can be explained by the horizontal movement of bicarbonate. Bicarbonate moves from above the nicotinamide ring towards the hydride and is placed in near-perpendicular position in relation to the $C4_N$–hydride bond (Figure 13, panels 5 and 6).

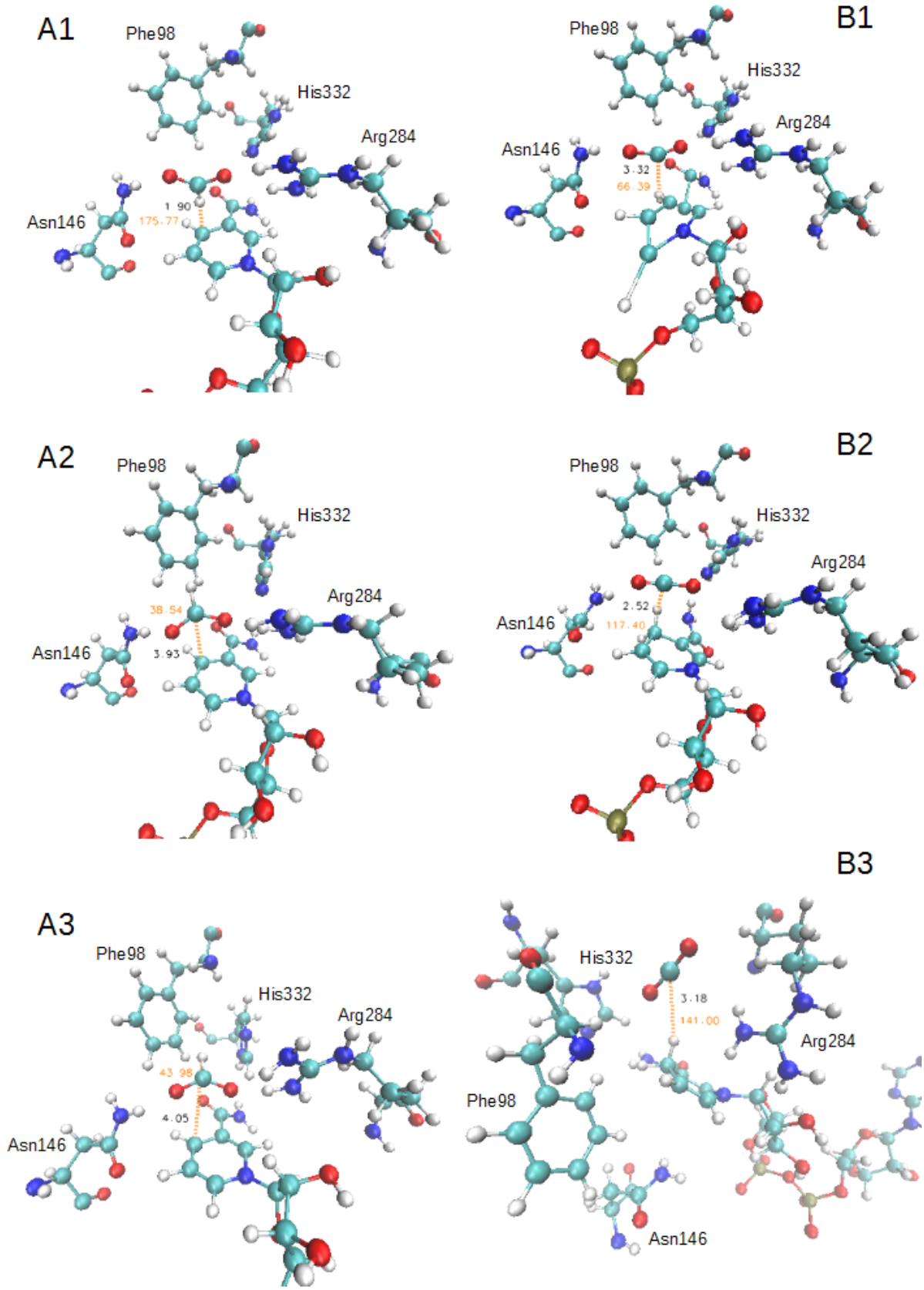


**Figure 12.** Conformations of formate (left panels; A) and $CO_2$ (right panels; B) during simulations. Acceptor carbon–hydride distances and angles are portrayed by black and orange numbers, respectively. Legend: (A1/B1) $t = 0$; (A2/B2) $t = 2$ ps (after minimization); (A3) $t = 2.70$ ns (final frame); (B3) $t = 4.05$ ns (final frame). Images made with VMD (Humphrey *et al.*, 1996).

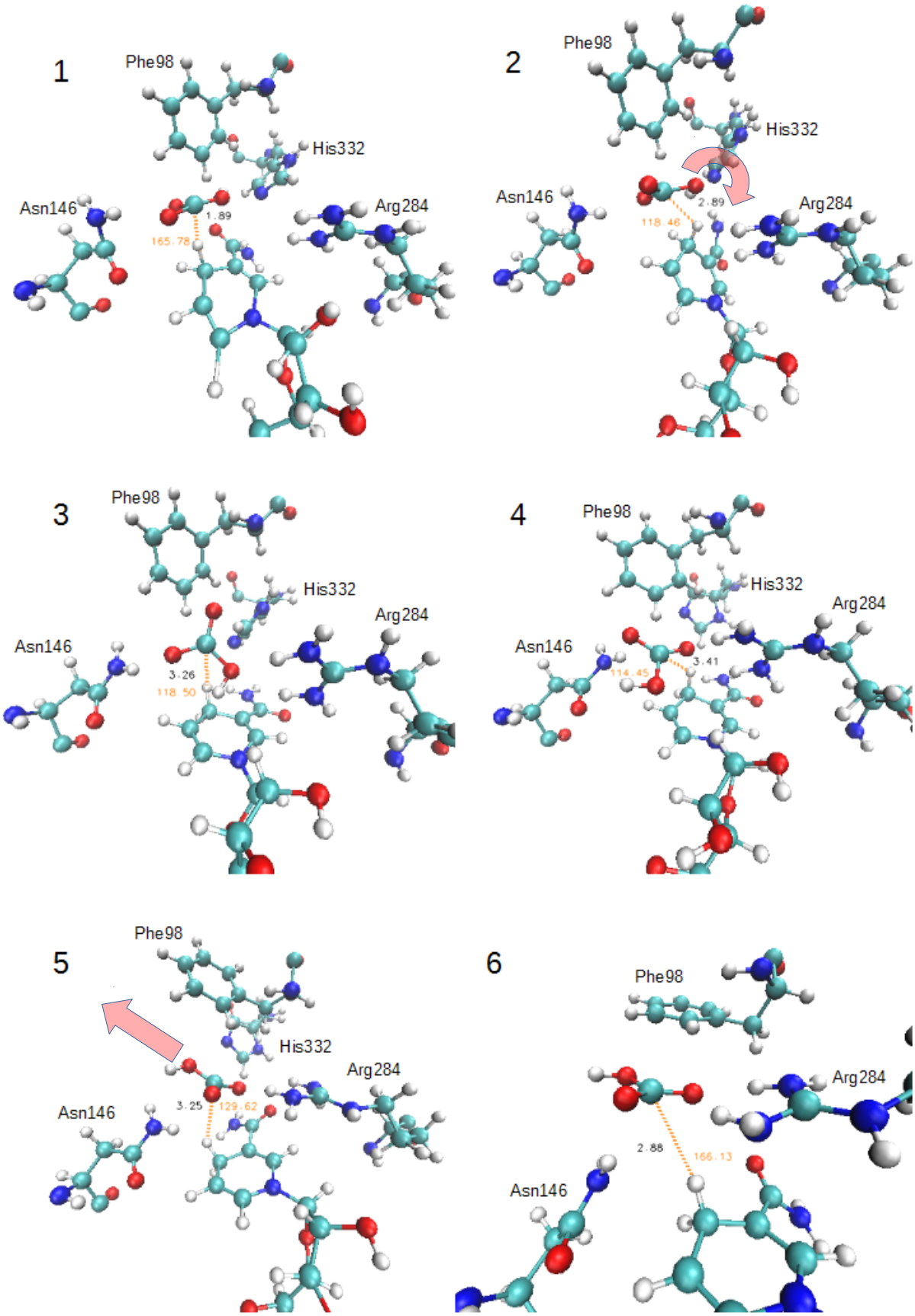


**Figure 13.** Conformation of bicarbonate during the simulation. Legend: (1) $t = 0$; (2) $t = 2$ ps, after minimization; (3) $t = 50$ ps, bicarbonate flips around over the nicotinamide ring; (4) $t = 2.43$ ns, bicarbonate drifts freely in the vicinity of the hydride; (5) $t = 2.87$ ns, slow migration towards the near attack conformation; (6) $t = 4.80$ ns, example of an optimal conformation for hydride transfer. Images made with VMD (Humphrey *et al*., 1996).

## 5.3 Proposed reaction mechanism for the reduction of bicarbonate

Based on the optimal conformation of bicarbonate in the active site of PseFDH, a novel model for the reduction of bicarbonate was developed (Figure 14). In the model, the hydride from NADH attacks the carbon atom of bicarbonate. Then, a short-lived transition state is formed. The overall reaction yields hydroxide and formate. In principle, formate could react further with the newly formed $NAD^+$. However, it is yet unclear if the hydroxide ion formed in the previous reaction could hinder the latter reaction.

### 5.3.1 Reaction mechanism in detail

The hydride transfer from NADH to bicarbonate is initiated when the near attack conformation is attained. In this model, the conditions for hydride transfer are assumed to be the same as for formate and $NAD^+$ defined by Torres *et al*. (1999). That is, the angle between the carbon atom of bicarbonate, the transferring hydride and the $C4_N$ does not necessarily have to be perpendicular at 180 degrees. Instead, a more tilted angle may be favored, as proposed by Wu & Houk (1987) who also stated that in hydride transfers between two planar molecules, the syn conformation is favored over anti. Both conformations are present in the simulation but the conditions of the near attack conformation are not fulfilled when the syn conformation is predominant (Figure 13, panel 5). The distance between the carbon atom of bicarbonate and the $C4_N$ does not descend to values under 3 Å until bicarbonate drifts to the anti conformation. It is possible that this is caused by amino acid residues Ile122, Asn146 and Arg284, which were evolved to direct formate towards the $C4_N$. Similar interactions with bicarbonate may lead to its placement in the anti conformation.

The short-lived and unstable transition state is formed when the hydride attacks the *p* orbital of the bicarbonate carbon which is double bonded to one of the oxygen atoms (Figure 14). To determine which oxygen atom has the negative charge and which is double bonded to carbon, one must take the charges of the environment into account. Logically, the oxygen atom which forms hydrogen bonds to positively charged Arg284 should be the negative one, because it would balance the differences in charges (see Figure 13, panel 6). Consequently, the oxygen atom facing Ile122 and Asn146 is involved in the double bond. The binding of the hydride to the carbon atom causes the double bond to break. The binding electrons of the double bond move to oxygen and as a result, the transition complex has two negatively charged oxygen

atoms. The hydroxyl group has remained unaffected until now when the repulsion between the negative oxygen atoms detach it from the molecule. Following the breakdown of the transition complex, formate and hydroxide are formed. Formate might react further to $CO_2$ but the fate of the hydroxide ion is unclear. In theory, the positively charged Arg284 could attract hydroxide and pull it into the substrate channel, discharging it from the active site.

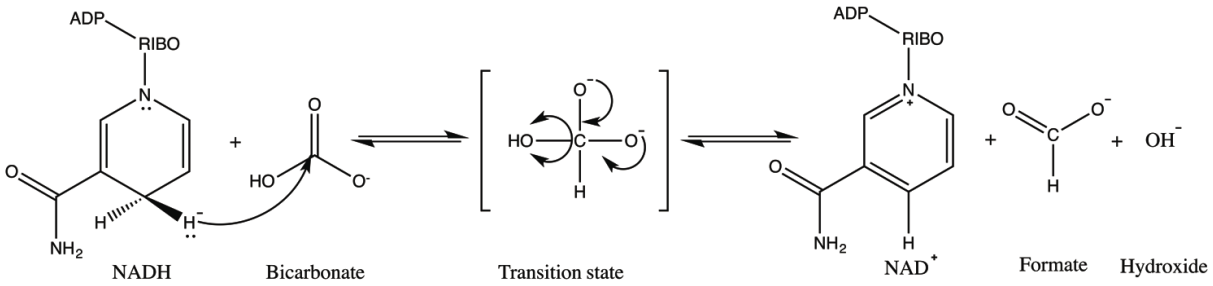


**Figure 14.** Proposed reaction mechanism for the reduction of bicarbonate yielding formate and hydroxide. Scheme adapted from Aslan *et al*. (2016) by the courtesy of Jarkko Valjakka (original author of the scheme).

## 5.3.2 Kinetic properties of the reaction

Aslan *et al*. (2016) have conducted a yet unpublished experimental study on the enzyme kinetics of formate dehydrogenases from *Chaetomium thermophilum* (CtFDH) and *Candida methylica* (CmFDH) using bicarbonate as substrate. Detailed results of these experiments can be reviewed in Tables 3 & 4 and Figure 15. Perhaps the most intriguing result was the fact that CtFDH was able to generate formate from bicarbonate. This was not the case for CmFDH which did not show appreciable bicarbonate–to–formate transformation rates. Due to high similarity between formate dehydrogenases, the results of this study could be applied to PseFDH with caution. However, it should be kept in mind that CtFDH is a fungal enzyme while both CmFDH and PseFDH originate from methylotrophic bacteria.

**Table 3.** Kinetic properties of CtFDH with different substrates. Data adapted from Aslan *et al*. (2016).

| **CtFDH** | $K_M$ (mM) | $k_{cat}$ ($s^{-1}$) | $k_{cat}$ / $K_M$ ($mM^{-1}s^{-1}$) |
|---|---|---|---|
| Formate | 3.30 ± 0.26 | 20.39 ± 0.46 | 6.18 |
| Bicarbonate | 0.36 ± 0.03 | 0.12 ± 0.00 | 0.34 |
| $NAD^+$ | 0.84 ± 0.08 | 33.40 ± 0.82 | 39.76 |
| NADH | 0.03 ± 0.00 | 0.22 ± 0.00 | 7.21 |

**Table 4.** Kinetic properties of CmFDH with different substrates. Data adapted from Aslan *et al*. (2016).

| **CmFDH** | $K_M$ (mM) | $k_{cat}$ ($s^{-1}$) | $k_{cat}$ / $K_M$ ($mM^{-1}s^{-1}$) |
|---|---|---|---|
| Formate | 7.01 ± 0.37 | 13.09 ± 0.26 | 1.87 |
| Bicarbonate | 0.78 ± 0.15 | 0.08 ± 0.00 | 0.10 |
| $NAD^+$ | 1.00 ± 0.12 | 31.69 ± 1.08 | 31.61 |
| NADH | 0.06 ± 0.00 | 0.11 ± 0.00 | 1.90 |

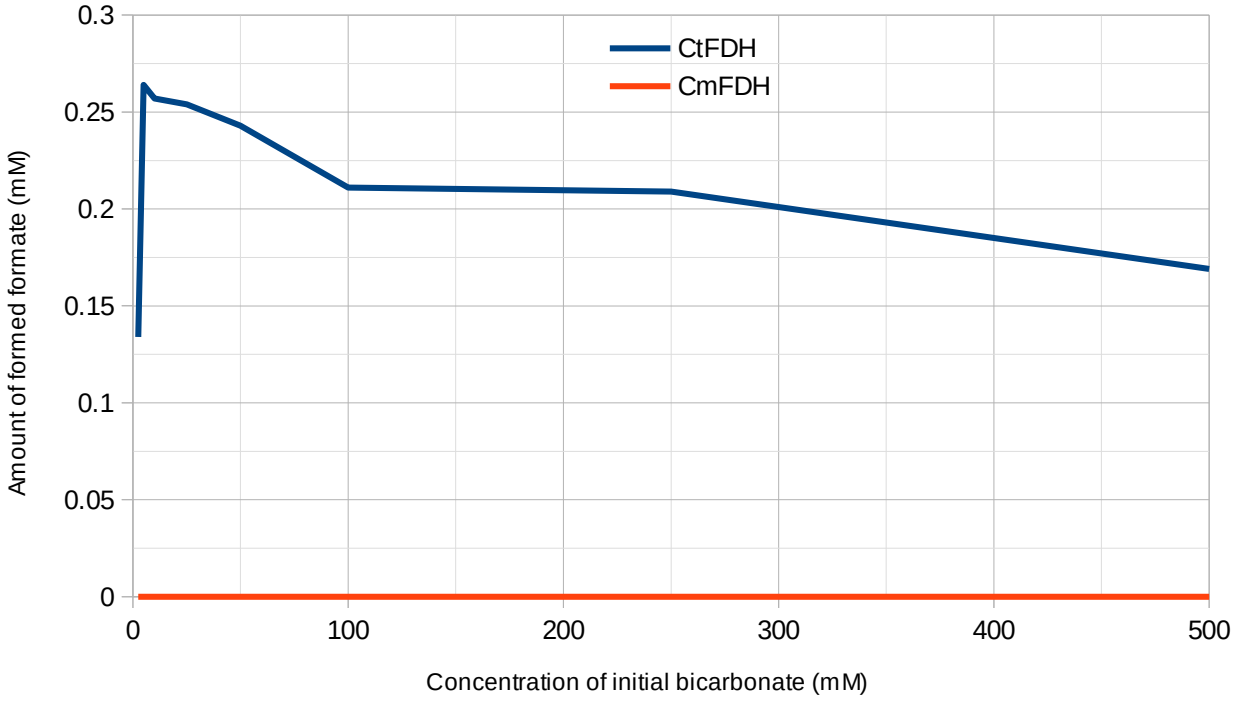


**Figure 15.** Effect of initial bicarbonate concentration on the activities of CtFDH and CmFDH. Data adapted from Aslan *et al*. (2016).

The experimental results are in line with the simulations performed for this thesis. Apparently, the enzyme is quickly saturated by bicarbonate, because the $K_M$ value is very low. The conversion from bicarbonate to formate and hydroxide may take up to several nanoseconds because bicarbonate needs time to reach the near attack conformation. In the simulation, the minimum time needed for proper positioning of bicarbonate was three nanoseconds. Before positioning itself to the near attack conformation, bicarbonate must gain access into the active site in the first place. The diffusion rate is likely another important contributor to the slow reaction speed. Overall, the results are promising, but further enzyme engineering is still needed for the bicarbonate conversion reaction to become profitable.

## 5.4 Future prospects

Based on the existing results, the binding interactions and reaction mechanism for bicarbonate in the active site of formate dehydrogenases could be studied further using quantum mechanics/molecular mechanics simulations. These more advanced simulations allow studying the hydride transfer on orbital level and simulating the effects of charges more accurately. Another important research subject is how bicarbonate travels to the active site.

Hypothetically, bicarbonate could use the very same substrate channel as formate does. It is though possible that the channel is too narrow for bicarbonate, and its only way to the active site is through the coenzyme channel.

Analysis of different mutations and their effects on the ability of formate dehydrogenases to utilize bicarbonate first *in silico* and then experimentally is a crucial step if more efficient systems for coenzyme regeneration or carbon fixation are needed. Native formate dehydrogenases are none too powerful for these purposes in large or even global scale.

# 6 Conclusion

The molecular dynamics simulations conducted in this thesis project have shown that bicarbonate can adapt a favorable conformation for hydride transfer in the active site of formate dehydrogenase from *Pseudomonas sp*., strain 101. The initial conformation which was constructed by optimizing the hydrogen bond interactions was found not to be optimal as bicarbonate settled in a conformation highly similar to that of formate during the course of the simulation. That is, Ile122, Asn146 and Arg284 form hydrogen bonds with the oxygen atoms of bicarbonate and keep the carbon atom slightly under 3 Å away from the hydride bound to NADH.

A plausible reaction mechanism for the hydride transfer from NADH to bicarbonate, yielding formate and hydroxide, was developed. The hydride must attack the *p* orbital of the carbon atom which forms a double bond to one of the oxygen atoms. Experimental data seems to support both the results of the computational analyses and the model of the reduction of bicarbonate. The orientation of bicarbonate towards the near attack conformation is a time consuming process and thus it is likely the rate-limiting step of the reaction.

A noteworthy observation about the research methods is the inaccuracy of classical molecular dynamics simulations under certain circumstances. Nevertheless, the classical MD method is still valuable when analyzing enzyme–substrate interactions *ab initio*. The preliminary results could then be used as a base for deeper analyses with quantum mechanics/molecular mechanics simulations or laboratory experiments.

# 7 References


Alekseeva AA, Savin SS, Tishkov VI. (2011). "$NAD^+$-dependent formate dehydrogenase from plants". *Acta Naturae*. **3** (4): 38–54.

Aslan AS, Turunen O, Valjakka J, Ruupunen J, Yıldırım D, Binay B. (2016). "*Chaetomium thermophilum* formate dehydrogenase as an example to illustrate the specific mechanism responsible for reduction of $CO_2$ to formic acid". Unpublished manuscript.

Brooks BR, Brooks CL 3rd, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M. (2009). "CHARMM: the biomolecular simulation program". *Journal of Computational Chemistry*. **30** (10): 1545–1614.

Brooks, BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus, M. (1983). "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations". *Journal of Computational Chemistry*. **4** (2): 187–217.

Egorov AM, Avilova TV, Dikov MM, Popov VO, Rodionov YV, Berezin IV. (1979). "NAD-dependent formate dehydrogenase from methylotrophic bacterium, strain 1. Purification and characterization". *European Journal of Biochemistry.* **99** (3): 569–576.

Galkin AG, Kutsenko AS, Bajulina NP, Esipova NG, Lamzin VS, Mesentsev AV, Shelukho DV, Tikhonova TV, Tishkov VI, Ustinnikova TB, Popov VO. (2002). "Site-directed mutagenesis of the essential arginine of the formate dehydrogenase active centre". *Biochimica Et Biophysica Acta.* **1594** (1): 136–149.

Guex N, Peitsch MC. (1997). "SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling". *Electrophoresis.* **18** (15): 2714–2723.

Humphrey W, Dalke A, Schulten K. (1996). "VMD: visual molecular dynamics". *Journal of Molecular Graphics.* **14** (1): 33–38.

Lamzin VS, Aleshin AE, Strokopytov BV, Yukhnevich MG, Popov VO, Harutyunyan EH, Wilson KS. (1992). "Crystal structure of NAD-dependent formate dehydrogenase". *European Journal of Biochemistry.* **206** (2): 441–452.

Lamzin VS, Dauter Z, Popov VO, Harutyunyan EH, Wilson KS. (1994). "High resolution structures of holo and apo formate dehydrogenase". *Journal of Molecular Biology.* **236** (3): 759–785.

Laskowski RA. (2001). "PDBsum: summaries and analyses of PDB structures". *Nucleic Acids Research*. **29** (1): 221–222.

Mondal B, Song J, Neese F, Ye S. (2015). "Bio-inspired mechanistic insights into $CO_2$ reduction". *Current Opinion in Chemical Biology*. **25**: 103–109.

Nilov DK, Shabalin IG, Popov VO, Svedas VK. (2012). "Molecular modeling of formate dehydrogenase: the formation of the Michaelis complex". *Journal of Biomolecular Structure & Dynamics.* **30** (2): 170–179.

Peacock D, Boulter D. (1970). "Kinetic studies of formate dehydrogenase". *Biochemical Journal.* **120** (4): 763–769.

Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. (2005). "Scalable molecular dynamics with NAMD". *Journal of Computational Chemistry.* **26** (16): 1781–1802.

Popov VO, Lamzin VS. (1994). "$NAD^+$-dependent formate dehydrogenase". *Biochemical Journal.* **301** (Pt 3): 625–643.

Reda T, Plugge CM, Abram NJ, Hirst J. (2008). "Reversible interconversion of carbon dioxide and formate by an electroactive enzyme". *Proc Natl Acad Sci U S A*. **105** (31): 10654–10658.

Sakai K, Hsieh BC, Maruyama A, Kitazumi Y, Shirai O, Kano K. (2015). "Interconversion between formate and hydrogen carbonate by tungsten-containing formate dehydrogenase-catalyzed mediated bioelectrocatalysis". *Sensing and Bio-Sensing Research*. **5**: 90–96.

Schiøtt B, Zheng Y, Bruice TC. (1998). "Theoretical investigation of the hydride transfer from formate to $NAD^+$ and the implications for the catalytic mechanism of formate dehydrogenase". *Journal of the American Chemical Society.* **120** (29): 7192–7200.

Senn HM, Thiel W. (2009). "QM/MM methods for biomolecular systems". *Angewandte Chemie*. **48** (7): 1198–1229.

Tishkov VI, Matorin AD, Rojkova AM, Fedorchuk VV, Savitsky PA, Dementieva LA, Lamzin VS, Mezentzev AV, Popov VO. (1996). "Site-directed mutagenesis of the formate dehydrogenase active centre: role of the His332-Gln313 pair in enzyme catalysis". *FEBS Letters*. **390** (1): 104–108.

Tishkov VI, Popov VO. (2004). "Catalytic mechanism and application of formate dehydrogenase". *Biochemistry-Russia.* **69** (11): 1252–1267.

Tishkov VI, Popov VO. (2006). "Protein engineering of formate dehydrogenase". *Biomolecular Engineering.* **23** (2-3): 89–110.

Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I, Mackerell AD Jr. (2010). "CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields". *Journal of Computational Chemistry.* **31** (4): 671–690.

Wu Y, Houk KN. (1987). "Theoretical transition structures for hydride transfer to methyleniminium ion from methylamine and dihydropyridine. On the nonlinearity of hydride transfers". *Journal of the American Chemical Society.* **109** (7): 2226–2227.

Wu Y, Lai DKW, Houk KN. (1995). "Transition structures of hydride transfer reactions of protonated pyridinium ion with 1,4-dihydropyridine and protonated nicotinamide with 1,4-

dihydronicotinamide". *Journal of the American Chemical Society.* **117** (14): 4100–4108.

Yu W, He X, Vanommeslaeghe K, MacKerell AD Jr. (2012). "Extension of the CHARMM General Force Field to sulfonyl-containing compounds and its utility in biomolecular simulations". *Journal of Computational Chemistry.* **33** (31): 2451–2468.

Özgün G, Karagüler NG, Turunen O, Turner NJ, Binay B. (2015). "Characterization of a new acidic NAD+-dependent formate dehydrogenase from thermophilic fungus Chaetomium thermophilum". *Journal of Molecular Catalysis B: Enzymatic.* **122** 212–217.

# Appendix

This section contains the topology files of ligands used in PSF file generation. The topology files were initially taken from the CGenFF (Vanommeslaeghe *et al*., 2010) and the atom names were modified to match those in the PDB files of the initial FDH–ligand structures.

### Formate

```
RESI FOR          -1.00 ! CHO2 formate, from acetate, sz & kevo
GROUP
ATOM  C1  CG2O3    0.52 !     H
ATOM  O2  OG2D2   -0.76 !     C
ATOM  O3  OG2D2   -0.76 !    / \\
ATOM  H4  HGR52    0.00 !  -O   O

BOND C1 H4
BOND C1 O2 C1 O3
IMPR C1 O2 O3 H4
IC  O3  O2 *C1 H4  0.00 0.00  180.0  0.0  0.0
IC  H4  O3 *C1 O2  0.00 0.00  180.0  0.0  0.0 !redundant definition needed to
enable seeding.

END
```

### Carbon dioxide

```
RESI CO2           0.00 ! CO2 Carbon Dioxide, John Straub
GROUP
ATOM C1   CG2O7    0.60
ATOM O1   OG2D5   -0.30
ATOM O2   OG2D5   -0.30
BOND O1  C1  O2  C1

END
```

### Bicarbonate

```
RESI BIC         -1.00 ! CHO3 bicarbonate, xxwy & kevo
ATOM C    CG2O6    0.69 ! H3    O1
ATOM H3   HGP1     0.43 ! |    /
ATOM O1   OG2D2   -0.76 ! O3--C   (-)
ATOM O2   OG2D2   -0.76 !      \
ATOM O3   OG311   -0.60 !       O2

BOND C  O1  C  O2  C  O3  O3 H3
IMPR C  O1 O2 O3
! seed = O1 C O2
IC O2  O1 *C O3   0.0000   0.00  180.00   0.00   0.0000
IC O1  C  O3 H3   0.0000   0.00    0.00   0.00   0.0000

END
```

## $NAD^+$

```
RESI NAD          -1.00  ! C21H26N7O14P2 oxidized nicotinamide adenine
dinucleotide, jjp1/adm jr.
                         ! atom names correspond to pdb nomenclature
                         ! checked for consistency with new NA params, adm jr.,
9/98
                         ! note that differences with respect to published results
exist
                         ! due to new NA params
GROUP
ATOM C4B  CG3C51    0.11  !                H61A  H62A
ATOM H4AX HGA1      0.09  !                   \  /
ATOM O4B  OG3C51   -0.40  !                   N6A
ATOM C1B  CG3C51    0.11  !                    |
ATOM H1AX HGA1      0.09  !                   C6A
GROUP                     !                  //  \
ATOM C5A  CG2RC0    0.28  !                 N1A  C5A--N7A\\
ATOM N7A  NG2R50   -0.71  !                  |    ||     C8A-H8A
ATOM C8A  CG2R53    0.34  !                 C2A  C4A--N9A/
ATOM H8A  HGR52     0.12  !                 / \\ /     \
ATOM N9A  NG2R51   -0.05  !              H2A  N3A       \
                         !                               \
ATOM N1A  NG2R62   -0.74  !                                \
ATOM C2A  CG2R64    0.50  !                                 \
ATOM H2A  HGR62     0.13  !    O1N     O1A    H5ASH4AX O4B  \
ATOM N3A  NG2R62   -0.75  !     |       |      |    \ /   \  \
ATOM C4A  CG2RC0    0.43  ! O5D PN-O3--PA-O5B -C5B -C4B    C1B
ATOM C6A  CG2R64    0.46  ! \   |       |      |     \     / \
                         !  \ O2N     O2A    H5AX C3B -C2B  H1AX
ATOM N6A  NG2S3    -0.77  !   \                     / \   / \
ATOM H61A HGP4      0.38  !    \                 O3B  H3AXO2B H2AX
ATOM H62A HGP4      0.38  !     \                  |       |
GROUP                     !      \               H3AT     H2AT
ATOM C2B  CG3C51    0.14  !       \                        H71N
ATOM H2AX HGA1      0.09  !        \                         \
ATOM O2B  OG311    -0.65  !         \                      H72N-N7N    H4N
ATOM H2AT HGP1      0.42  !          \                             \     |
GROUP                     !           \                           C7N  C4N
ATOM C3B  CG3C51    0.14  !            \                         // \  / \\
ATOM H3AX HGA1      0.09  !             \                     O7N  C3N  C5N-H5N
ATOM O3B  OG311    -0.65  !              \                          ||    |
ATOM H3AT HGP1      0.42  !               \ H5NS H4NXO4D       H2N-C2N C6N-H6N
GROUP                     !                \ |    \ /   \            \+//
ATOM C5B  CG321    -0.08  !                 C5D --C4D    C1D ---------N1N
ATOM H5AX HGA2      0.09  !                  |     \     / \
ATOM H5AS HGA2      0.09  !                 H5NX   C3D -C2D H1NX
ATOM PA   PG1       1.50  !                       / \   / \
ATOM O1A  OG2P1    -0.82  !                   O3D  H3NXO2D H2NX
ATOM O2A  OG2P1    -0.82  !                    |       |
ATOM O5B  OG303    -0.62  !                   H3NT    H2NT
ATOM O3   OG304    -0.68
ATOM PN   PG1       1.50
ATOM O1N  OG2P1    -0.82
ATOM O2N  OG2P1    -0.82
ATOM O5D  OG303    -0.62
ATOM C5D  CG321    -0.08
ATOM H5NS HGA2      0.09
ATOM H5NX HGA2      0.09
GROUP
ATOM C2D  CG3C51    0.14
ATOM H2NX HGA1      0.09
ATOM O2D  OG311    -0.65
ATOM H2NT HGP1      0.42
GROUP
```

```
ATOM C3D  CG3C51    0.14
ATOM H3NX HGA1      0.09
ATOM O3D  OG311    -0.65
ATOM H3NT HGP1      0.42
GROUP
ATOM C1D  CG3C53    0.11
ATOM H1NX HGA1      0.09
ATOM C4D  CG3C51    0.11
ATOM H4NX HGA1      0.09
ATOM O4D  OG3C51   -0.40
GROUP
ATOM N1N  NG2R61   -0.07
ATOM C6N  CG2R62    0.16
ATOM H6N  HGR63     0.19
ATOM C5N  CG2R62   -0.10
ATOM H5N  HGR63     0.16
ATOM C4N  CG2R62   -0.05
ATOM H4N  HGR63     0.16
ATOM C3N  CG2R62    0.05
ATOM C2N  CG2R62    0.18
ATOM H2N  HGR63     0.16
ATOM C7N  CG2O1     0.68
ATOM O7N  OG2D1    -0.40
ATOM N7N  NG2S2    -0.82
ATOM H71N HGP1      0.34  ! trans to amide O
ATOM H72N HGP1      0.36  ! cis to amide O

BOND N1A  C2A       N3A  C4A       C5A  C6A
BOND C6A  N6A       C5A  N7A       C8A  N9A
BOND N9A  C4A       C2A  H2A       N6A  H61A      N6A  H62A      C8A  H8A
DOUBLE C6A  N1A       C2A  N3A       C4A  C5A      N7A  C8A
BOND N9A  C1B       C1B  C2B       C2B  C3B       C3B  C4B       C4B  O4B
BOND O4B  C1B       C1B  H1AX      C2B  O2B       O2B  H2AT      C2B  H2AX
BOND C3B  H3AX      C3B  O3B       O3B  H3AT      C4B  H4AX      C4B  C5B
BOND C5B  H5AS      C5B  H5AX      C5B  O5B       O5B  PA        PA   O1A
BOND PA   O2A       PA   O3        O3   PN        PN   O1N       PN   O2N
BOND PN   O5D       O5D  C5D       C5D  H5NS      C5D  H5NX      C5D  C4D
BOND C4D  O4D       O4D  C1D       C1D  C2D       C2D  C3D       C3D  C4D
BOND C1D  H1NX      C2D  H2NX      C2D  O2D       O2D  H2NT      C3D  H3NX
BOND C3D  O3D       O3D  H3NT      C4D  H4NX      C1D  N1N       N1N  C2N
BOND C3N  C4N       C5N  C6N
BOND C2N  H2N       C3N  C7N       C7N  O7N       C7N  N7N       N7N  H71N
BOND N7N  H72N      C4N  H4N       C5N  H5N       C6N  H6N
DOUBLE C2N  C3N       C4N  C5N       C6N  N1N
IMPR C6A   C5A   N1A   N6A
IMPR N6A   H62A  H61A  C6A
IMPR C7N   C3N   N7N   O7N
DONO H61A N6A
DONO H62A N6A
DONO H2AT O2B
DONO H3AT O3B
ACCE N1A
ACCE N3A
ACCE N7A
ACCE O4B
ACCE O2B
ACCE O3B
ACCE O5B
ACCE O1A  PA
ACCE O2A  PA
ACCE O3
ACCE O1N  PN
ACCE O2N  PN
ACCE O5D
ACCE O4D
```

```
ACCE O3D
ACCE O2D
ACCE O7N
DONO H2NT O2D
DONO H3NT O3D
DONO H71N N7N
DONO H72N N7N
IC PA     O3     PN     O5D    1.4863   65.28 -169.00   98.59   1.5977
IC C5B    O5B    PA     O3     1.4232  127.31 -165.10  103.27   1.4863
IC C5B    O5B    PA     O2A    1.4232  127.31   73.33  111.48   1.4836
IC H5AS   C5B    C4B    C3B    0.9935  120.00  -58.20  111.58   1.6942
IC PA     O3     PN     O1N    1.4863   65.28   78.51  108.97   1.4756
IC PA     O3     PN     O2N    1.4863   65.28  -54.29  112.88   1.4636
IC PA     O5B    C5B    C4B    1.5901  127.31  121.29  111.63   1.5491
IC O1A    PA     O5B    C5B    1.5901  127.31  121.29  127.31   1.4232
IC O5B    C5B    C4B    C3B    1.4232  111.63  -58.20  111.58   1.6942
IC C5B    C4B    C3B    O3B    1.5491  111.58  128.42  114.19   1.4337
IC H3AT   O3B    C3B    C4B    0.9671   98.77  147.40  114.19   1.6942
IC O4B    C3B    *C4A   C5B    1.8868  112.95 -118.10  111.58   1.5491
IC C2B    C4B    *C3B   O3B    1.5097   93.22 -117.72  114.19   1.4337
IC C4B    C3B    C2B    C1B    1.6942   93.22  -12.93  117.82   1.5415
IC C3B    C2B    C1B    N9A    1.5097  117.82  135.56  115.01   1.4847
IC O4B    C1B    N9A    C4A    1.3646   95.44  -90.72  125.96   1.4013
IC C1B    C4A    *N9A   C8A    1.4847  125.96 -176.52  105.34   1.3777
IC C4A    N9A    C8A    N7A    1.4013  105.34   -0.07  114.01   1.3282
IC C8A    N9A    C4A    C5A    1.3777  105.34    0.11  105.06   1.3782
IC C8A    N7A    C5A    C6A    1.3282  103.21 -179.92  130.78   1.4146
IC N7A    C5A    C6A    N1A    1.3814  130.78 -179.94  117.85   1.3482
IC C5A    C6A    N1A    C2A    1.4146  117.85   -0.14  118.87   1.3300
IC N9A    C5A    *C4A   N3A    1.4013  105.06 -179.67  126.14   1.3648
IC C5A    N1A    *C6A   N6A    1.4146  117.85  179.89  119.69   1.3419
IC N1A    C6A    N6A    H61A   1.3482  119.69   -0.39  116.60   0.9912
IC H61A   C6A    *N6A   H62A   0.9912  116.60 -179.02  116.94   0.9978
IC C5A    N1A    *C6A   N6A    1.4146  117.85  179.89  119.69   1.3419
IC N1A    C6A    N6A    H61A   1.3482  119.69   -0.39  116.60   0.9912
IC H61A   C6A    *N6A   H62A   0.9912  116.60 -179.02  116.94   0.9978
IC N9A    N7A    *C8A   H8A    1.3777  114.01 -179.37  126.34   1.0962
IC N1A    N3A    *C2A   H2A    1.3300  129.42  179.97  114.82   1.0928
IC C1B    C3B    *C2A   O2B    1.5415  117.82 -145.06  114.13   1.4294
IC H2AX   O2B    C2B    C3B    2.0386   31.12  -93.63  114.13   1.5097
IC H2AT   O2B    C2B    C3B    0.9953   99.36  -93.63  114.13   1.5097
IC O4B    C2B    *C1B   H1AX   1.3646  120.77 -114.99  109.65   1.1105
IC C1B    C3B    *C2A   H2AX   1.5415  117.82  108.23   86.44   1.0999
IC C2B    C4B    *C3B   H3AX   1.5097   93.22  117.28  111.94   1.1110
IC C3B    O4B    *C4A   H4AX   1.6942  112.95 -140.26   57.99   1.0000
IC C4B    O5B    *C5A   H5AX   1.5491  111.63 -123.75  111.23   1.1111
IC C4B    O5B    *C5A   H5AX   1.5491  111.63 -123.75  111.23   1.1111
IC C5D    O5D    PN     O2N    1.4451  128.40  -49.72  108.83   1.4636
IC H5NS   C5D    O5D    PN     1.1110  109.50  115.00  128.40   1.5977
IC H5NX   C5D    O5D    PN     1.1110  109.50 -115.00  128.40   1.5977
IC PN     O5D    C5D    C4D    1.5977  128.40    0.00  110.10   1.5160
IC O5D    C5D    C4D    C3D    1.4451  110.10    0.00  108.50   1.5160
IC C5D    C4D    C3D    C2D    1.5160  108.50    0.00  111.00   1.5160
IC C4D    C3D    C2D    C1D    1.5160  111.00    0.00  105.50   1.5270
IC C3D    C2D    C1D    O4D    1.5160  105.50    0.00  105.00   1.4100
IC C2D    C1D    O4D    C4D    1.5270  105.00    0.00  117.86   1.4712
IC O2D    C2D    C1D    O4D    1.4200  110.10  180.00  105.00   1.4100
IC H2NT   O2D    C2D    C1D    0.9600  106.00  180.00  110.10   1.5270
IC O4D    C2D    *C1D   H1NX   1.4100  105.00 -115.00  110.10   1.1110
IC C1D    C3D    *C2N   H2NX   1.5270  105.50  115.00  110.10   1.1110
IC C2D    C4D    *C3N   H3NX   1.5160  111.00  115.00  110.10   1.1110
IC C3D    O4D    *C4N   H4NX   1.5160  100.64 -115.00  107.24   1.1110
IC C4D    O5D    *C5N   H5NX   1.5160  110.10 -115.00  109.50   1.1110
IC C4D    O5D    *C5N   H5NS   1.5160  110.10  115.00  109.50   1.1110
IC C3D    C2D    C1D    N1N    1.5160  105.50    0.00  113.70   1.4800
```

```
IC O3D    C3D    C2D    C1D    1.4200  110.10  180.00  105.50   1.5270
IC H3NT   O3D    C3D    C2D    0.9600  106.00  180.00  110.10   1.5160
IC C2D    C1D    N1N    C2N    1.5270  113.70    0.00  121.70   1.3150
IC C1D    N1N    C2N    C3N    1.4800  121.70    0.00  122.00   1.3500
IC N1N    C2N    C3N    C4N    1.3150  122.00    0.00  118.00   1.3600
IC C2N    C3N    C4N    C5N    1.3500  118.00    0.00  118.00   1.3600
IC C3N    C4N    C5N    C6N    1.3600  118.00    0.00  118.00   1.3500
IC C4N    C5N    C6N    N1N    1.3600  118.00    0.00  124.51   1.2199
IC C5N    C6N    N1N    C2N    1.3500  124.51    0.00  119.49   1.3150
IC N1N    C2N    C3N    C7N    1.3150  122.00    0.00  131.80   1.4800
IC C2N    C3N    C7N    O7N    1.3500  131.80    0.00  118.50   1.2300
IC C2N    C3N    C7N    N7N    1.3500  131.80    0.00  113.00   1.3600
IC O7N    C7N    N7N    H71N   1.2300  120.00  180.00  120.00   1.0000
IC O7N    C7N    N7N    H72N   1.2300  120.00    0.00  120.00   1.0000
IC C2N    C3N    C4N    H4N    1.3500  118.00    0.00  121.00   1.0900
IC C3N    C4N    C5N    H5N    1.3600  118.00    0.00  119.00   1.0900
IC C4N    C5N    C6N    H6N    1.3600  118.00    0.00  120.50   1.0900
IC C6N    N1N    C2N    H2N    1.2199  119.49    0.00  117.50   1.0900
PATCH FIRST NONE LAST NONE

END
```

## NADH

```
RESI NAD           -2.00  ! C21H27N7O14P2 reduced nicotinamide adenine
dinucleotide, jjp1/adm jr.
                          ! atom names correspond to pdb nomenclature
                          ! checked for consistency with new NA params, adm jr.,
9/98
                          ! note that differences with respect to published results
exist
                          ! due to new NA params

GROUP
ATOM C4B  CG3C51    0.11  !                H61A  H62A
ATOM H4AX HGA1      0.09  !                   \  /
ATOM O4B  OG3C51   -0.40  !                   N6A
ATOM C1B  CG3C51    0.11  !                    |
ATOM H1AX HGA1      0.09  !                   C6A
GROUP                     !                  //  \
ATOM C5A  CG2RC0    0.28  !                 N1A  C5A--N7A\\
ATOM N7A  NG2R50   -0.71  !                  |    ||     C8A-H8A
ATOM C8A  CG2R53    0.34  !                 C2A  C4A--N9A/
ATOM H8A  HGR52     0.12  !                 / \\ /     \
ATOM N9A  NG2R51   -0.05  !              H2A  N3A       \
                          !                              \
ATOM N1A  NG2R62   -0.74  !                               \
ATOM C2A  CG2R64    0.50  !                                \
ATOM H2A  HGR62     0.13  !    O1N     O1A    H5ASH4AX O4B  \
ATOM N3A  NG2R62   -0.75  !     |       |      |    \ /   \  \
ATOM C4A  CG2RC0    0.43  ! O5D PN-O3--PA-O5B -C5B -C4B    C1B
ATOM C6A  CG2R64    0.46  ! \   |       |      |     \     / \
                          !  \ O2N     O2A    H5AX C3B -C2B  H1AX
ATOM N6A  NG2S3    -0.77  !   \                      / \   / \
ATOM H61A HGP4      0.38  !    \                 O3B  H3AXO2B H2AX
ATOM H62A HGP4      0.38  !     \                  |       |
GROUP                     !      \               H3AT     H2AT
ATOM C2B  CG3C51    0.14  !       \                         H71N
ATOM H2AX HGA1      0.09  !        \                          \
ATOM O2B  OG311    -0.65  !         \                      H72N-N7N  H4N H4N2
ATOM H2AT HGP1      0.42  !          \                            \    \ /
GROUP                     !           \                           C7N  C4N
ATOM C3B  CG3C51    0.14  !            \                          / \  /  \
ATOM H3AX HGA1      0.09  !             \                      O7N  C3N  C5N-H5N
ATOM O3B  OG311    -0.65  !              \                           ||   ||
ATOM H3AT HGP1      0.42  !               \ H5NS H4NXO4D       H2N-C2N C6N-H6N
GROUP                     !                \ |    \ /   \            \  /
ATOM C5B  CG321    -0.08  !                 C5D --C4D    C1D ---------N1N
ATOM H5AX HGA2      0.09  !                  |     \     / \
ATOM H5AS HGA2      0.09  !                 H5NX   C3D -C2D H1NX
ATOM PA   PG1       1.50  !                       / \   / \
ATOM O1A  OG2P1    -0.82  !                   O3D  H3NXO2NXH2NX
ATOM O2A  OG2P1    -0.82  !                    |       |
ATOM O5B  OG303    -0.62  !                   H3NT    H2NT
ATOM O3   OG304    -0.68
ATOM PN   PG1       1.50
ATOM O1N  OG2P1    -0.82
ATOM O2N  OG2P1    -0.82
ATOM O5D  OG303    -0.62
ATOM C5D  CG321    -0.08
ATOM H5NS HGA2      0.09
ATOM H5NX HGA2      0.09
GROUP
ATOM C2D  CG3C51    0.14
ATOM H2NX HGA1      0.09
ATOM O2NX OG311    -0.65
ATOM H2NT HGP1      0.42
```

```
GROUP
ATOM C3D  CG3C51    0.14
ATOM H3NX HGA1      0.09
ATOM O3D  OG311    -0.65
ATOM H3NT HGP1      0.42
GROUP
ATOM C1D  CG3C51    0.11
ATOM H1NX HGA1      0.09
ATOM C4D  CG3C51    0.11
ATOM H4NX HGA1      0.09
ATOM O4D  OG3C51   -0.40
GROUP
ATOM N1N  NG301    -0.27 !N2
ATOM C6N  CG2D1O   -0.06 !C3
ATOM H6N  HGA4      0.17 !H4
ATOM C5N  CG2D1    -0.18 !C5
ATOM H5N  HGA4      0.14 !H6
ATOM C4N  CG321    -0.28 !C7
ATOM H4N  HGA2      0.09 !H8
ATOM H4N2 HGA2      0.09 !H17
ATOM C3N  CG2DC1    0.36 !C9
ATOM C2N  CG2D1O   -0.10 !C10
ATOM H2N  HGA4      0.14 !H11
ATOM C7N  CG2O1     0.55 !C12
ATOM O7N  OG2D1    -0.51 !O13
ATOM N7N  NG2S2    -0.72 !N14
ATOM H71N HGP1      0.26 !H15  ! trans to amide O
ATOM H72N HGP1      0.32 !H16  ! cis to amide O

BOND N1A  C2A       N3A  C4A       C5A  C6A
BOND C6A  N6A       C5A  N7A       C8A  N9A
BOND N9A  C4A       C2A  H2A       N6A  H61A      N6A  H62A      C8A  H8A
DOUBLE  C6A  N1A      C2A  N3A      C4A  C5A      N7A  C8A
BOND N9A  C1B       C1B  C2B       C2B  C3B       C3B  C4B       C4B  O4B
BOND O4B  C1B       C1B  H1AX      C2B  O2B       O2B  H2AT      C2B  H2AX
BOND C3B  H3AX      C3B  O3B       O3B  H3AT      C4B  H4AX      C4B  C5B
BOND C5B  H5AS      C5B  H5AX      C5B  O5B       O5B  PA        PA   O1A
BOND PA   O2A       PA   O3        O3   PN        PN   O1N       PN   O2N
BOND PN   O5D       O5D  C5D       C5D  H5NS      C5D  H5NX      C5D  C4D
BOND C4D  O4D       O4D  C1D       C1D  C2D       C2D  C3D       C3D  C4D
BOND C1D  H1NX      C2D  H2NX      C2D  O2NX      O2NX H2NT      C3D  H3NX
BOND C3D  O3D       O3D  H3NT      C4D  H4NX      C1D  N1N       N1N  C2N
BOND C3N  C4N       C4N  C5N       C6N  N1N
BOND C2N  H2N       C3N  C7N       C7N  O7N       C7N  N7N       N7N  H71N
BOND N7N  H72N      C4N  H4N       C4N  H4N2      C5N  H5N       C6N  H6N
DOUBLE  C2N  C3N      C5N  C6N
IMPR C6A   C5A   N1A   N6A
IMPR N6A   H62A  H61A  C6A
IMPR C6N   C5N   N1N   H6N
IMPR C2N   C3N   N1N   H2N
IMPR C7N   C3N   N7N   O7N
DONO H61A N6A
DONO H62A N6A
DONO H2AT O2B
DONO H3AT O3B
ACCE N1A
ACCE N3A
ACCE N7A
ACCE O4B
ACCE O2B
ACCE O3B
ACCE O5B
ACCE O1A  PA
ACCE O2A  PA
ACCE O3
```

```
ACCE O1N  PN
ACCE O2N  PN
ACCE O5D
ACCE O4D
ACCE O3D
ACCE O2NX
ACCE O7N
DONO H2NT O2NX
DONO H3NT O3D
DONO H71N N7N
DONO H72N N7N
! IC table was beyond repair ==> replaced (kevo)
IC PA     O3     PN     O5D    1.6011   92.88  167.64   96.12   1.6600
IC PA     O3     PN     O1N    1.6011   92.88   62.12  118.93   1.5360
IC PA     O3     PN     O2N    1.6011   92.88  -85.34  112.47   1.5263
IC PN     O3     PA     O5B    1.6074   92.88  169.27   93.16   1.6636
IC O3     O5B    *PA    O1A    1.6011   93.16 -115.15  102.32   1.5255
IC O3     O5B    *PA    O2A    1.6011   93.16  120.82  100.53   1.5378
IC O3     PA     O5B    C5B    1.6011   93.16  -65.25  117.89   1.4434
IC PA     O5B    C5B    C4B    1.6636  117.89 -111.08  111.72   1.5437
IC C4B    O5B    *C5B   H5AX   1.5437  111.72  121.91  113.03   1.1186
IC C4B    O5B    *C5B   H5AS   1.5437  111.72 -120.24  108.95   1.1118
IC O5B    C5B    C4B    C3B    1.4434  111.72   32.23  115.30   1.5403
IC C5B    C4B    C3B    O3B    1.5437  115.30   82.44  113.07   1.4340
IC H3AT   O3B    C3B    C4B    0.9956  102.19  -78.33  113.07   1.5403
IC O4B    C3B    *C4B   C5B    1.4436  105.34 -122.02  115.30   1.5437
IC C2B    C4B    *C3B   O3B    1.5312  103.02 -124.10  113.07   1.4340
IC C4B    C3B    C2B    C1B    1.5403  103.02   34.98  101.18   1.5414
IC C3B    C2B    C1B    N9A    1.5312  101.18   88.32  111.78   1.4609
IC O4B    C1B    N9A    C4A    1.4336  105.95 -149.10  127.35   1.3682
IC C1B    C4A    *N9A   C8A    1.4609  127.35  179.89  106.85   1.3818
IC C4A    N9A    C8A    N7A    1.3682  106.85    0.34  112.92   1.3338
IC C8A    N9A    C4A    C5A    1.3818  106.85   -0.12  105.55   1.3970
IC C8A    N7A    C5A    C6A    1.3338  103.64 -179.39  132.41   1.4083
IC N7A    C5A    C6A    N1A    1.3924  132.41  179.90  118.87   1.3616
IC C5A    C6A    N1A    C2A    1.4083  118.87   -0.29  119.53   1.3626
IC N9A    C5A    *C4A   N3A    1.3682  105.55 -179.81  125.91   1.3450
IC C5A    N1A    *C6A   N6A    1.4083  118.87  179.60  117.11   1.3448
IC N1A    C6A    N6A    H61A   1.3616  117.11 -176.96  119.87   0.9980
IC H61A   C6A    *N6A   H62A   0.9980  119.87  174.83  117.95   0.9930
IC N9A    N7A    *C8A   H8A    1.3818  112.92 -179.84  125.33   1.0978
IC N1A    N3A    *C2A   H2A    1.3626  125.37 -179.78  117.21   1.0932
IC O4B    C2B    *C1B   H1AX   1.4336  107.49 -118.51  111.70   1.1012
IC C1B    C3B    *C2B   H2AX   1.5414  101.18  120.60  111.56   1.0994
IC C2B    C4B    *C3B   H3AX   1.5312  103.02  114.75  109.97   1.1020
IC C3B    O4B    *C4B   H4AX   1.5403  105.34 -116.63  105.95   1.1009
IC C5D    O5D    PN     O2N    1.4353  125.45 -110.33  103.18   1.5263
IC PN     O5D    C5D    C4D    1.6600  125.45 -100.01  114.36   1.5573
IC O5D    C5D    C4D    C3D    1.4353  114.36   42.03  115.36   1.5403
IC C5D    C4D    C3D    C2D    1.5573  115.36  -91.03  101.62   1.5376
IC C4D    C3D    C2D    C1D    1.5403  101.62  -39.21  100.72   1.5399
IC C3D    C2D    C1D    O4D    1.5376  100.72   35.40  105.65   1.4395
IC O2NX   C2D    C1D    O4D    1.4392  115.62  158.69  105.65   1.4395
IC H2NT   O2NX   C2D    C1D    0.9686  102.30  157.15  115.62   1.5399
IC O4D    C2D    *C1D   H1NX   1.4395  105.65 -114.11  109.56   1.1015
IC C1D    C3D    *C2N   H2NX   2.4605    9.08  131.65   26.20   3.7695
IC C2D    C4D    *C3N   H3NX   4.7247   22.53   -7.48   17.99   6.8615
IC C3D    O4D    *C4N   H4NX   6.4634   19.31  -56.73    2.93   7.0769
IC C4D    O5D    *C5N   H5NX   5.1412   29.29   35.30   14.10   6.2151
IC C4D    O5D    *C5N   H5NS   5.1412   29.29   55.33   24.87   4.9161
IC C3D    C2D    C1D    N1N    1.5376  100.72  153.61  113.06   1.4871
IC O3D    C3D    C2D    C1D    1.4390  111.69   79.56  100.72   1.5399
IC H3NT   O3D    C3D    C2D    0.9633  102.31   55.04  111.69   1.5376
IC C2D    C1D    N1N    C2N    1.5399  113.06  115.33  119.55   1.3595
IC C1D    N1N    C2N    C3N    1.4871  119.55  179.60  123.83   1.3763
```

```
IC N1N    C2N    C3N    C4N    1.3595  123.83    5.03  118.94   1.5428
IC C2N    C3N    C4N    C5N    1.3763  118.94   -6.70  111.53   1.5189
IC C3N    C4N    C5N    C6N    1.5428  111.53    5.08  122.32   1.3550
IC N1N    C2N    C3N    C7N    1.3595  123.83 -158.70  118.33   1.5212
IC C2N    C3N    C7N    O7N    1.3763  118.33 -160.59  120.47   1.2354
IC C2N    C3N    C7N    N7N    1.3763  118.33   20.47  119.04   1.3683
IC O7N    C7N    N7N    H71N   1.2354  120.48 -176.29  120.56   0.9982
IC O7N    C7N    N7N    H72N   1.2354  120.48    0.52  116.91   0.9963
IC C2N    C3N    C4N    H4N    1.3763  118.94 -136.51  113.66   1.1068
IC C2N    C3N    C4N    H4N2   1.3763  118.94  108.33  105.96   1.1115
IC C3N    C4N    C5N    H5N    1.5428  111.53 -149.76  116.93   1.0907
IC C4N    C5N    C6N    H6N    1.5189  122.32  179.97  119.29   1.0969
IC C6N    N1N    C2N    H2N    1.3589  121.05  177.95  116.69   1.0900
IC C1B    C3B    *C2B   O2B    1.5414  101.18 -118.90  110.89   1.4319
IC C3B    C2B    O2B    H2AT   1.5312  110.89  -28.83  100.98   0.9715
PATCH FIRST NONE LAST NONE

END
```